U.S. DEPARTMENT OF COMMERCE
National Technical Information Service

PB-296 838

Statistical Analysis of TSP and Meteorological Data in EPA Region 6

Technology Service Corp, Santa Monica, CA

Prepared for

Environmental Protection Agency, Dallas, TX Surveillance and Analysis Div

Jan 79

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO.	2.	3. RECIPIENT'S ACCESSION NO. PB296838
4. TITLE AND SUBTITLE STATISTICAL ANALYSIS OF TSP AND METEOROLOGICAL DATA IN EPA REGION 6		5. REPORT DATE January 1979
		6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) John Trijonis Yuji Horie Dennis Bicker		8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS Technology Service Corporation 2811 Wilshire Boulevard Santa Monica, CA 90403		10. PROGRAM ELEMENT NO.
		11. CONTRACT/GRANT NO. 68-02-2828
12. SPONSORING AGENCY NAME AND ADDRESS Surveillance and Analysis Division Environmental Protection Agency - Region 6 1201 Elm Street Dallas, Texas 75202		13. TYPE OF REPORT AND PERIOD COVERED
		14. SPONSORING AGENCY CODE 906/00

15. SUPPLEMENTARY NOTES

16. ABSTRACT

This report conducts statistical analyses of TSP and meteorological data at 25 locations in EPA Region 6. The primary goal of the study is to determine if the relationship between TSP and meteorology can be used to classify sites as to the types of sources causing nonattainment.

It is found that decision-tree analysis is a very effective means of characterizing the relationship between TSP and meteorology, and that the TSP-meteorology relationship provides useful insights concerning the sources of high TSP levels. The results indicate distinct geographical patterns in causes of nonattainment within Region 6; specifically, wind-blown dust tends to be a major cause in and near the Texas Panhandle and a very minor cause in Arkansas and (probably) Louisiana. Despite a fair level of success with the method, however, we conclude that the TSP-meteorology relationship does not provide results that are definitive or quantitative enough for it to be used as the sole support for regulatory policy concerning causes of nonattainment.

In addition to decision-tree analyses of the TSP-meteorology relationship, several other statistical studies are conducted herein. These include analyses of data quality, investigations of cumulative frequency distributions using log-normal and exponential distribution functions, computations of various statistical parameters for TSP data, regression studies of the TSP-meteorology relationship, nomogram analysis of the TSP-meteorology relationship, regression analysis of time series data, and meteorological normalization of time series data.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
Air Pollution Total Suspended Particulates Meteorology Data Analysis Fugitive Dust	Hi-Vol Data Decision-Tree Analysis Emission Source Identification Data Quality Control Wind-Blown Dust	

18. DISTRIBUTION STATEMENT RELEASE UNLIMITED	19. SECURITY CLASS *(This Report)* UNCLASSIFIED	21. NO. OF PAGES 238
	20. SECURITY CLASS *(This page)* UNCLASSIFIED	22. PRICE A11 - A01

EPA Form 2220-1 (9-73)

PB 296838

EPA 906/9-79-005

STATISTICAL ANALYSIS
OF TSP AND METEOROLOGICAL DATA
IN EPA REGION 6

Prepared by

John Trijonis
Yuji Horie
Dennis Bicker
Technology Service Corporation
Santa Monica, California

for the

United States Environmental Protection Agency
1201 Elm Street, Suite 2800
Dallas, Texas 75270

Contract #68-02-2828

Miguel Flores, Project Officer

REPRODUCED BY
NATIONAL TECHNICAL
INFORMATION SERVICE
U. S. DEPARTMENT OF COMMERCE
SPRINGFIELD, VA. 22161

This document is available to the public through the National Technical Information Service, Springfield, Virginia 22151.

This report has been reviewed by EPA Region 6 and approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the Environemntal Protection Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

ABSTRACT

This report conducts statistical analyses of TSP and meteorological data at 25 locations in EPA Region 6. The primary goal of the study is to determine if the relationship between TSP and meteorology can be used to classify sites as to the types of sources causing nonattainment.

It is found that decision-tree analysis is a very effective means of characterizing the relationship between TSP and meteorology, and that the TSP- meteorology relationship provides useful insights concerning the sources of high TSP levels. The results indicate distinct geographical patterns in causes of nonattainment within Region 6; specifically, wind-blown dust tends to be a major cause in and near the Texas Panhandle and a very minor cause in Arkansas and (probably) Louisiana. Despite a fair level of success with the method, however, we conclude that the TSP-meteorology relationship does not provide results that are definitive or quantitative enough for it to be used as the sole support for regulatory policy concerning causes of nonattainment.

In addition to decision-tree analyses of the TSP-meteorology relationship, several other statistical studies are conducted herein. These include analyses of data quality, investigations of cumulative frequency distributions using log-normal and exponential distribution functions, computations of various statistical parameters for TSP data, regression studies of the TSP-meteorology relationship, nomogram analysis of the TSP-meteorology relationship, regression analysis of time series data, and meteorological normalization of time series data.

TABLE OF CONTENTS

Section Page

1.0 INTRODUCTION. 1

1.1 Summary of Results and Conclusions. 2
1.2 Recommendations 6

2.0 SELECTION OF MONITORING SITES AND DESCRIPTION OF SITE CHARACTERISTICS 8

2.1 Selection of Study Locations. 8
2.2 TSP Monitoring Site Characteristics 11
2.2.1 Monitoring Site Types 11
2.2.2 TSP Emission Density. 15
2.2.3 Climatic Characteristics. 15

3.0 QUALITY CHECK OF TSP DATA 26

3.1 Outlier Detection Using Upper and Lower Thresholds. 26
3.2 Outlier Detection Using Percentile Concentration Plots. . . . 26
3.3 Outlier Detection From Site-By-Site Comparisons 28
3.4 Comfirmation of Outliers. 33

4.0 DETERMINATION OF STATISTICAL PARAMETERS FOR THE TSP DATA. 45

4.1 Summary Statistics of TSP Data. 45
4.2 Cumulative Frequency Distributions. 49

5.0 DECISION-TREE ANALYSIS OF TSP AND METEOROLOGICAL DATA 51

5.1 The AID Decision-Tree Program 52
5.2 Application of "AID" to TSP and Meteorological Data 53
5.2.1 Meteorological Variables. 53
5.2.2 Results of AID. 56
5.3 Comparison of AID Results to Regression Results 60

6.0 INTERPRETATION OF THE DECISION-TREES AND VALIDATION OF THE RESULTS 64

6.1 Use of the TSP-Meteorology Relationship for Identifying Source Types. 64
6.2 Interpretation of the AID Decision-Trees. 68
6.3 Validation of the Results for Texas Locations 69
6.4 Evaluation of the Method. 74

Section Page

7.0 REGIONAL PATTERNS IN THE CONTRIBUTION OF WIND-BLOWN DUST TO TSP CONCENTRATIONS. 78

7.1 An Alternative Analysis of the TSP-Meteorology Relationship. 78
7.2 Patterns in Causes of High TSP Levels 85

8.0 TIME SERIES OF TSP DATA. 91

8.1 Time Series Plots and Regression Analysis 91
8.2 Meteorological Adjustment of Ambient TSP Data 93

9.0 REFERENCES 98

APPENDICES

A. Cumulative Frequency Distributions of TSP Data 101

B. AID Decision-Trees for the 25 Study Locations. 153

C. Interpretations of AID Decision-Trees 179

D. Time Series Plots of TSP Data 203

FIGURES

Number Page

1 Map of study region showing boundaries of AQCRs and states. . . 10

2 Location of 25 TSP monitoring stations shown in the regional map of AQCRs 12

3 Combinations of TSP monitoring stations and upper air weather stations. 13

4 Location of 25 TSP monitoring stations shown in the Regional Map of State Climatic Divisions. 18

5 Mean annual precipitation, based on the period 1931-55. 19

6 Annual number of days with 0.01 inch or more of precipitation . 20

7 Mean annual relative humidity in percent. 21

8 Annual surface wind roses 22

9 Prevailing surface wind direction and annual mean wind speed. . 24

10 Annual total heating degree days. 25

11 Log-normal distribution plot of TSP data at Jonesboro, AR . . . 29

12 Exponential distribution plot of TSP data at Corpus Christi, Texas 30

13 Log-normal distribution plot of TSP data at Stuttgart, AR . . . 31

14 Exponential distribution plot of TSP data at Stuttgart, AR. . . 32

15a Corrected log-normal distribution plot of TSP data at Las Cruces, New Mexico. 41

15b Corrected exponential distribution plot of TSP data at Las Cruces, New Mexico. 42

16a Corrected log-normal distribution plot of TSP data at Tulsa, Oklahoma 43

16b Corrected exponential distirbution plot of TSP data at Tulsa, Oklahoma 44

17 Exponential distribution plot of TSP data at Bethany, OK. . . . 48

18 Hypothetical example of nonmonotonic relationship that would be discerned by AID 54

19 AID decision-tree for Forrest City, Arkansas. 57

20 Geographical distribution of sites classified according to causes of high TSP levels. 87

21 Annual average number of hours with wind-blown dust during the 1950's 89

Number Page

22 Annual percent frequency of hours with blowing dust observations, 1940-1970 90

23 Time series plot for Harlingen, Texas 92

24 Time series at Harlingen TX, stratified by meteorological class 94

25 Meteorological adjustment of annual mean TSP concentrations at Harlingen, Texas. 96

TABLES

Number — Page

1 Sites Selected for EPA Region 6 TSP Study. 9
2 Summary of Site Characteristics 14
3 Summary of Site Classifications 15
4 Annual TSP Emissions by County. 16
5 Outliers of TSP Data in Region 6 27
6 Outliers Detected By Site-By-Site Comparisons of TSP Data . . . 34
7 Check of TSP Data Flagged as Outliers by Telephone Inquiry. . . 38
8 Summary Statistics of TSP Data In The Study Area. 46
9 Sensitivity of Attainment Status to Yearly Sampling Frequency . 49
10 Summary of Meteorological Variables 55
11 Summary of Meteorological Variables Selected by AID 59
12 Percent Variance Explained by "AID" Compared to Percent Variance Explained by Multiple Regressions Against All 18 Meteorological Variables. 61
13 Significant Meteorological Variables Identified by AID and by Multiple Regressions 63
14 Interpretation of Decision-Tree for Forrest City, AR. 70
15 Summary of Conclusions From the Interpretation of the AID Decision-Trees. 71
16 Summary of TACB Conclusions Concerning the 7 Texas Locations 73
17 Comparison of TACB Conclusions to the Conclusions Obtained by Analyzing the TSP/Meteorology Decision-Trees 75
18 Nomogram for Classifying Days of High TSP According to Probability of Wind-Blown Dust. 79
19 Distribution of All Sampling Days Into Five Classes for Distinguishing Probability of Wind-Blown Dust 80
20 Distribution of High TSP Levels as to Probability of Wind-Blown Dust According to the Nomogram 82
21 Comparison of the Results from the AID and Nomogram Analyses Concerning Contributions of Wind-Blown Dust to TSP. 84
22 Three-Class Categorization of Sites With Respect to Wind-Blown Dust Contribution. 86

1.0 INTRODUCTION AND SUMMARY

A significant number of locations in EPA Region 6 presently violate the National Ambient Air Quality Standards (NAAQS) for total suspended particulate matter (TSP). In order to formulate realistic control strategies for TSP, the causes of these violations must be determined. Specifically, there is a need to ascertain the relative contributions of wind-blown dust, anthropogenic fugitive-dust sources, and conventional particulate sources.

Three recent studies, conducted by the Texas Air Control Board, (Price et al. 1977), GCA Corporation (Record et al. 1976), and TRW Inc. (Richard et al. 1977), have attempted to identify the origins of suspended particulates causing violations of the NAAQS. For some locations these studies have met with success using advanced analytical methods (such as microscopy, X-ray fluorescence, and tracer elements), detailed site surveys, and/or analysis of TSP, meteorological, and emission data. However, none of these studies has provided a simple, reliable, widely applicable method of determining particulate origins.

The principal objective of this investigation is to determine whether modern data-analytic techniques, applied to TSP and meteorological data, can be used to categorize TSP monitoring sites as to cause of nonattainment. Specifically, decision-tree analysis is used to characterize the relationship between TSP and meteorology; the TSP-meteorology relationship is then interpreted to identify the types of sources causing high TSP concentrations. This study also applies other statistical techniques to investigate data quality, frequency distributions, and time series of TSP data.

This report is organized in 8 chapters. The present chapter provides an introduction and a summary of conclusions and recommendations. Chapter 2 describes the 25 Region 6 locations to be studied and documents relevant site characteristics. Chapter 3 conducts a quality check of the three years of TSP data (1973-1975) used in the study. Chapter 4 discusses the statistical distributions of TSP data for the various locations. Chapter 5 describes the application of the "AID" decision-tree program to TSP and meteorological data and compares the decision-tree results to regression results.

Chapter 6 interprets the decision-trees, identifies sources of high TSP levels, and validates the results for selected sites. Chapter 7 supplements the decision-tree analysis with a nomogram analysis and describes regional patterns in the contribution of wind-blown dust to TSP. Chapter 8 analyzes the TSP data as time series and illustrates how the decision-tree output can be used to normalize the time series for meteorology.

1.1 SUMMARY OF RESULTS AND CONCLUSIONS

The paragraphs below summarize the findings and conclusions reached in this study. For convenient reference, the summary is organized according to the order of the chapters.

Selection of Monitoring Sites and Description of Site Characteristics (Chapter 2)

- Twenty-five nonattainment monitoring sites in 11 Air Quality Control Regions within EPA Region 6 are selected for study. These sites are grouped so that sufficient weather data for the analysis can be provided by 10 surface weather stations and 6 upper air weather stations.
- The 25 study locations represent a wide variety of site characteristics. The characteristics documented herein include geographic location, distance to weather stations, SAROAD site type (local environment), county emission density of conventional TSP sources, climatology, and elevation of monitor above ground.

Quality Check of TSP Data (Chapter 3)

- Prior to analyzing aerometric data, it is important to verify data quality. Potentially erroneous data are identified using three methods: (1) selection of values outside upper and lower thresholds; (2) outlier detection by examining the upper tail of the cumulative frequency distribution (using both log-probability and exponential paper); and (3) comparisons of TSP values at different sites on the same date. The verification procedures for potentially erroneous data include checking the data by contacting local monitoring agencies, examining meteorological conditions on the sampling dates, and (when possible) comparing the data to readings at nearby sites on the same sampling date.

- Several zero TSP readings were invalidated; one zero value at an Albuquerque site was a transcription error, and ten zero values at a Tulsa site represented missing data. Two high TSP readings (one at Las Cruces and one at a Tulsa site) were invalidated by checking with the local monitoring agencies.

Determination of Statistical Parameters for the TSP Data (Chapter 4)

- At most study locations a lognormal distribution fits well to the main body of the TSP data. Deviations from the lognormal fit often occur, however, in the extreme upper tail of the cumulative frequency distribution, indicating a separate lognormal function should be used to characterize extreme values. Exponential distributions do not fit the main body and lower tail of the data very well, but they usually correspond well to the upper tail of the data. Because of this property and because plots on exponential paper emphasize resolution in the upper tail of the data, the exponential distribution is used herein to calculate expected yearly highest (or second highest) TSP concentrations.

- Summary statistics prepared for TSP concentrations at each study location include the arithmetic mean, geometric mean, geometric standard deviation, actual yearly highest/second-highest values, and expected yearly highest/second-highest values. Among the study locations, the geometric means range from 38 to 160 $\mu g/m^3$; the expected yearly maxima for 60 samples per year range from 150 to 850 $\mu g/m^3$; and the expected yearly maxima for 365 samples per year range from 200 to 1450 $\mu g/m^3$.

Decision-Tree Analysis of TSP and Meteorological Data (Chapter 5)

- The AID decision-tree program developed at the University of Michigan is used to characterize the relationship between TSP concentrations and meteorology. The AID program uses a very general form of the least-squares principle to divide the data into meteorological classes that best explain the variance in daily TSP levels. One advantage of the decision-tree approach is that the meteorological classes can be readily interpreted as to their physical meaning.

- The AID program is applied at each of the 25 study locations using 18 meteorological variables. The significant meteorological variables, and the form of their relationship to TSP, vary from site-to-site. Meteorological variables that are usually important include wind speed parameters, wind direction parameters, mixing heights, precipitation parameters, temperatures, month, and relative humidity.

- The AID decision-tree program is a very effective tool for analyzing the relationship between TSP and meteorology. The percent variance explained in TSP by AID (using a 10-group split) averages 51% among the 25 study locations. Multiple-linear regression and multiple-log linear regression using all 18 meteorological variables explain only 36% and 39% of the variance, respectively. As expected, the variables that are statistically significant in the regressions are usually also significant in the AID analysis.

Interpretation of the Decision-Trees and Validation of the Results (Chapter 6)

- The relationship between TSP and meteorology can provide insight into whether high TSP levels are caused by wind-blown dust (dust raised by the action of winds), man-made sources (dust raised by anthropogenic activities), or conventional sources (particulate sources contained in conventional emission inventories, such as particulates in exhaust plumes). Useful insights can be gained from several types of weather parameters: precipitation, surface wind speed, wind direction, wind variability, vertically averaged wind speed, mixing height, relative humidity, temperature, visibility, and month.

- Interpretation of the AID decision-trees yields a qualitative description of the types of sources causing high TSP levels at all 25 study locations. For the seven Texas locations included in this study, the results of the AID analysis agree fairly well with the conclusions reached in a much more comprehensive study performed by the Texas Air Control Board.

- Despite the reasonable level of success in applying and interpreting the AID analysis, we conclude that it is not feasible to use the TSP-meteorology relationship alone as a means of classifying sites as to cause of nonattainment. The results of the AID analyses are only of a qualitative nature, and they often involve ambiquities which prevent definitive classification of some sites. Also, there is a need for finer resolution in the types of sources that are identified; in particular, it would be useful to distinguish what part of wind-blown dust represents regional dust storms (which can reasonably be considered uncontrollable) and what part represents localized wind-blown dust from agricultural fields, storage piles, construction sites, etc. (which may be amenable to some control).

Regional Patterns in the Contribution of Wind-Blown Dust to TSP Concentrations (Chapter 7)

- To supplement the AID analysis, a two dimensional nomogram (involving wind speed and precipitation) is used to sort the data and to classify sites as to the probability that wind-blown dust causes the observed violations of the short-term TSP standards. The nomogram results tend to implicate wind-blown dust more strongly than the AID results. Two plausible reasons for this discrepancy are discussed.

- By considering both the AID and nomogram analyses, it is possible to form a 3-class categorization of sites: Class I - sites where wind-blown dust appears to be only a minor contributor to high TSP levels; Class II - sites where wind-blown dust appears to be a significant but not the major contributor to high TSP levels, or sites where there is large uncertainty; and Class III - sites where wind-blown dust appears to be the major contributor to high TSP levels. Four of the 25 study locations fall in Class I, and six fall in Class III. Partly because of uncertainties in the results, most sites are categorized as Class II.

- Consideration of our three-class categorization and of other results published in the literature reveals distinct geographical patterns in the contribution of wind-blown dust to TSP. We venture the following gross generalizations: In the Texas Panhandle or within approximately 150 miles of the Texas Panhandle, wind-blown dust tends to be a significant contributor, and often the dominant contributor, to high TSP concentrations. In Arkansas (and probably Louisiana) wind-blown dust tends to be a minor, often negligible contributor to high TSP concentrations. At other Region 6 locations, wind-blown dust may or may not be a significant source of high TSP levels.

- No obvious correlations appear between wind-blown dust contributions and SAROAD site type, county emission density, or elevation of monitor above ground. A basic climatological pattern exists (wind-blown dust is most significant in dry and windy areas); climatology is, of course, closely related to the geographical patterns in wind-blown dust contributions.

Time Series of TSP Data (Chapter 8)

- Time series plots and time series regressions identify several sites with significant upward or downward trends in TSP and several sites where temporary source activity, such as construction, may have been important. It is difficult to arrive at definitive conclusions, however, because we are dealing with only 3 years of data for an intermittently sampled pollutant. The small number of samples and the large variance produced by meteorological fluctuations hinder the interpretations of the time series.

- Meteorological classes defined by the AID analysis can be very useful in adjusting the time series for meteorological fluctuations and biases. This adjustment can be performed either by stratifying the data into meteorological classes or by conducting a quantitative meteorological normalization of air quality statistics. Again, however, problems are introduced by the small sample sizes which result from only 3 years of data for an intermittently sampled pollutant.

1.2 RECOMMENDATIONS

The analyses of TSP and meteorological data in EPA Region 6 lead to the following recommendations:

- Studies of TSP data should be preceded by data quality checks to eliminate erroneous values. Simple but useful methods of identifying potentially erroneous data include flagging values outside threshold limits and examining the cumulative frequency distribution. Although the log-normal function tends to fit the main body of TSP data fairly well, we recommend that the exponential distribution be used to examine the upper tail of the data. The exponential distribution usually fits well to the upper tail of the data, and plots on exponential paper emphasize resolution in the higher percentiles of the cumulative frequency distribution.

- Compliance monitoring for TSP is almost always conducted with intermittent sampling, e.g. 60 days per year sampling. Statistical methods should be applied to estimate the second-highest concentration for the entire year from data sets that are incomplete due to intermittent sampling and/or missing values. The exponential distribution function would likely be a useful tool in performing this estimation.

- The relationship between TSP and meteorology can provide useful insights as to the types of sources causing nonattaniment of TSP standards. In order to characterize the TSP-meteorology relationship, we recommend the use of decision-tree analysis. Decision-tree analysis involves less restrictive assumptions, produces more easily interpreted output, and explains more variance in TSP concentrations than multiple linear regression. Decision-tree analysis also appears well suited to handle the problems of few data points and many independent variables.

- We do not recommend that the TSP-meteorology relationship be used alone to classify sites as to cause of nonattainment for the purpose of regulatory planning. Further information, such as some of the factors included in the Texas Air Control Board study, should be considered to arrive at a more quantitative, accurate, and detailed assessment of causes for nonattainment. The only exception might be locations where the TSP-meteorology relationship points almost exclusively toward wind-blown dust.

- Geographical location within Region 6 should be considered in analyzing sites as to cause of nonattainment. Sites in and near the Texas Panhandle are likely to have wind-blown dust as the major source of high TSP levels. Sites in Arkansas (and probably Louisiana) are likely to have only minor contributions from wind-blown dust.

- A comprehensive study of TSP time series data would be useful to determine where control strategies are improving air quality, where source growth is worsening air quality, and where temporary source activities (such as construction) have produced temporarily high TSP levels. Meteorological adjustment of air quality trends should be performed as part of time series analysis of TSP data. Sites included in such a study should have at least five years of fairly complete (∿ 50 samples per year) data. It might be useful to select a subgroup of the sites included in this report for a pilot study of time series data.

2.0 SELECTION OF MONITORING SITES AND DESCRIPTION OF SITE CHARACTERISTICS

Twenty-five TSP monitoring sites in eleven Air Quality Control Regions (AQCRs) within EPA Region 6 are included in this study. The present chapter discusses the rationale for selecting the sites and documents the characteristics of the monitor locations.

2.1 SELECTION OF STUDY LOCATIONS

The choice of TSP monitoring sites for the study involved the following criteria (in descending order of importance):

1. Each site should be a nonattainment site for the annual geometric mean standard (75 μg/m^3) and/or the primary 24-hour standard (260 μg/m^3) for at least one year from 1973 to 1975.
2. The TSP data for the sites should be fairly complete for the years 1973 to 1975 (in most cases at least 100 samples during the three year period).
3. The sites should represent several geographical areas in EPA Region 6, a variety of climatic conditions, and a variety of local environments (e.g. rural vs. suburban vs. urban).
4. The sites should be located in large-scale spatial clusters to minimize the number of meteorological stations required.
5. The sites should be at varied distances from surface and upper air meteorological stations.

Based on these criteria, and with the aid of EPA Region 6 personnel who were familiar with the monitoring locations (Flores 1977) and regional climatology (Hall 1977), 25 sites in 11 AQCRs were selected for the analysis. Table 1 presents summary data for these sites, including AQCR, state, county, city, SAROAD site code, nonattainment status, and amount of TSP data in 1973, 74, and 75. Table 1 also lists the surface weather station and upper air weather station selected for use with each TSP site.

The study region encompassing the 25 sites consists of four states: Texas, New Mexico, Oklahoma, and Arkansas.* Figure 1 illustrates the study

---

*Note that EPA Region 6 also includes one other state, Louisiana.

TABLE 1. SITES SELECTED FOR EPA REGION 6 TSP STUDY

REGION		SITE				SURFACE WEATHER	UPPER AIR DATA	NON-ATTAINMENT		AMOUNT OF TSP DATA
AQCR	STATE	CITY		SITE CODE	COUNTY	NCC SITE	NCC SITE	AGM	PRIMARY 24 HR	73-74-75
211	TEXAS	LUBBOCK	(LU)	453340 001F01	LUBBOCK	LUBBOCK	AMARILLO	A	P	50-32-32
213	TEXAS	HARLINGEN	(HA)	452320 001F01	TEXAS	BROWNSVILLE	BROWNSVILLE	A	P	59-50-47
		McALLEN	(MC)	453390 001F01	HIDALGO	BROWNSVILLE	BROWNSVILLE	A		51-46-40
		SAN BENITO	(SB)	454600 001F01	CAMERON	BROWNSVILLE	BROWNSVILLE	A	P	49-42-49
214	TEXAS	CORPUS CHRISTI	(CC)	451150 001F01	NUECES	CORPUS CHRISTI	BROWNSVILLE	A		44-40-50
		CORPUS CHRISTI	(CC)	451150 004F02	NUECES	CORPUS CHRISTI	BROWNSVILLE	A	P	31-48-49
		CORPUS CHRISTI	(CC)	451150 019F01	NUECES	CORPUS CHRISTI	BROWNSVILLE	A		16-46-49
017	ARKANSAS	SPRINGDALE	(SP)	042480 001F01	WASHINGTON	FORT SMITH	MONET	A	P	58-58-54
	OKLAHOMA	SEQUOYAH CO.	(SE)	372760 471F01	SEQUOYAH	FORT SMITH	MONET		P	63-41-37
020	ARKANSAS	FORREST CITY	(FC)	040900 001F01	ST. FRANCIS	LITTLE ROCK	LITTLE ROCK	A	P	58-59-55
		HELENA	(HE)	041080 001F01	PHILLIPS	LITTLE ROCK	LITTLE ROCK	A		59-61-54
		JONESBORO	(JO)	041320 001F01	CRAIGHEAD	LITTLE ROCK	LITTLE ROCK	A		56-57-59
		STUTTGART	(ST)	042540 001F01	ARKANSAS	LITTLE ROCK	LITTLE ROCK	A	P	60-61-53
		STUTTGART	(ST)	042540 005F05	ARKANSAS	LITTLE ROCK	LITTLE ROCK	A	P	30-52-49
187	OKLAHOMA	ROGER MILLS	(RM)	372600 900F01	ROGER MILLS	OKLAHOMA CITY	AMARILLO		P	33-37-43
184	OKLAHOMA	BETHANY	(BE)	370260 014F01	OKLAHOMA	OKLAHOMA CITY	MONET	A		96-82-45
		OKLAHOMA CITY	(OK)	372200 002F01	OKLAHOMA	OKLAHOMA CITY	MONET		P	98-67-0
186	OKLAHOMA	TULSA	(TU)	373000 110F01	TULSA	TULSA	MONET		P	84-112-47
		TULSA	(TU)	373000 112F01	TULSA	TULSA	MONET	A	P	109-110-41
		MUSKOGEE	(MU)	371980 162F01	MUSKOGEE	TULSA	MONET		P	94-91-80
152	NEW MEX.	ALBUQUERQUE	(AL)	320040 003H01	BERNALILLO	ALBUQUERQUE	ALBUQUERQUE	A	P	0-45-54
		ALBUQUERQUE	(AL)	320040 008H02	BERNALILLO	ALBUQUERQUE	ALBUQUERQUE	A		4-42-53
153	NEW MEX.	DONA ANA	(DA)	320340 001F02	DONA ANA	EL PASO	EL PASO	A	P	9-86-47
		LAS CRUCES	(LC)	320580 001F01	DONA ANA	EL PASO	EL PASO	A	P	0-75-44
154	NEW MEX.	RATON	(RA)	320900 001F01	COLFAX	CLAYTON	AMARILLO		P	0-43-46

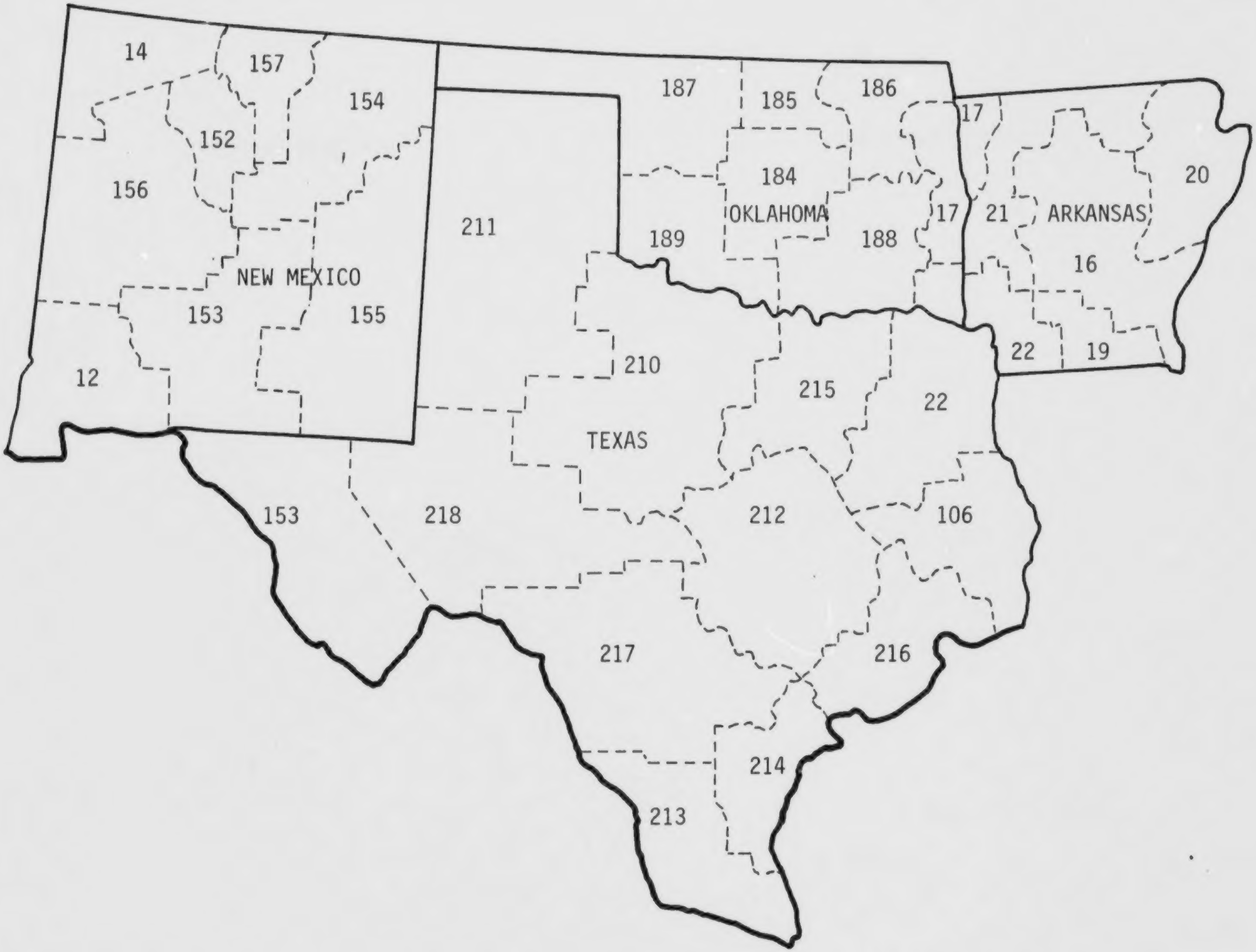

14
157
154
152
156
NEW MEXICO
153
155
12
211
187
185
186
184
OKLAHOMA
189
188
17
17
21
ARKANSAS
20
16
22
19
210
215
22
TEXAS
153
218
212
106
217
216
214
213

Figure 1. Map of study region showing boundaries of AQCRs and states.

region and delineates the geographical boundary of each AQCR. The locations of the 25 sites are illustrated in Figure 2. This figure also shows the 10 surface weather sites that will supply meteorological data for the analysis. Figure 3 illustrates the 6 upper air weather stations that will be used in the study.

2.2 TSP MONITORING SITE CHARACTERISTICS

For use in interpreting our later results concerning causes of high TSP levels, the following characteristics were documented for each of the 25 sites:

- SAROAD site type
- County emission density (conventional particulate sources)
- Climatology
- Elevation of monitor above ground

Table 2 provides a summary of these site characteristics. As to emission and climatic characteristics, the ranges of values among the 25 sites are divided into five sub-ranges: high, moderately high, moderate, moderately low, and low. These ratings are relative to normal conditions within EPA Region 6.

2.2.1 Monitoring Site Types

The SAROAD site types include two classifications--center city (CC), suburban (S) or rural (R); and commercial (C), industrial (I), residential (R), or agricultural (A). EPA Region 6 personnel (Flores 1977) arrived at the site types listed in Table 2 by reviewing site surveys and contacting local monitoring agencies. These classifications agree with those in the "Directory of Air Quality Monitoring Sites Active in 1974" (EPA 1976) with the following exceptions: the 1974 Directory lists no site type for Harlingen (001F01), lists Corpus Christi (001F01) as S/C, lists Oklahoma City (001F02) as CC/C, and lists Las Cruces (001F01) as S/C.

Table 3 summarizes the site types included in the 25 locations under study. It is difficult to obtain a wide variety of site types because the great preponderance of monitoring locations is in the center city or suburban and commercial or residential categories.

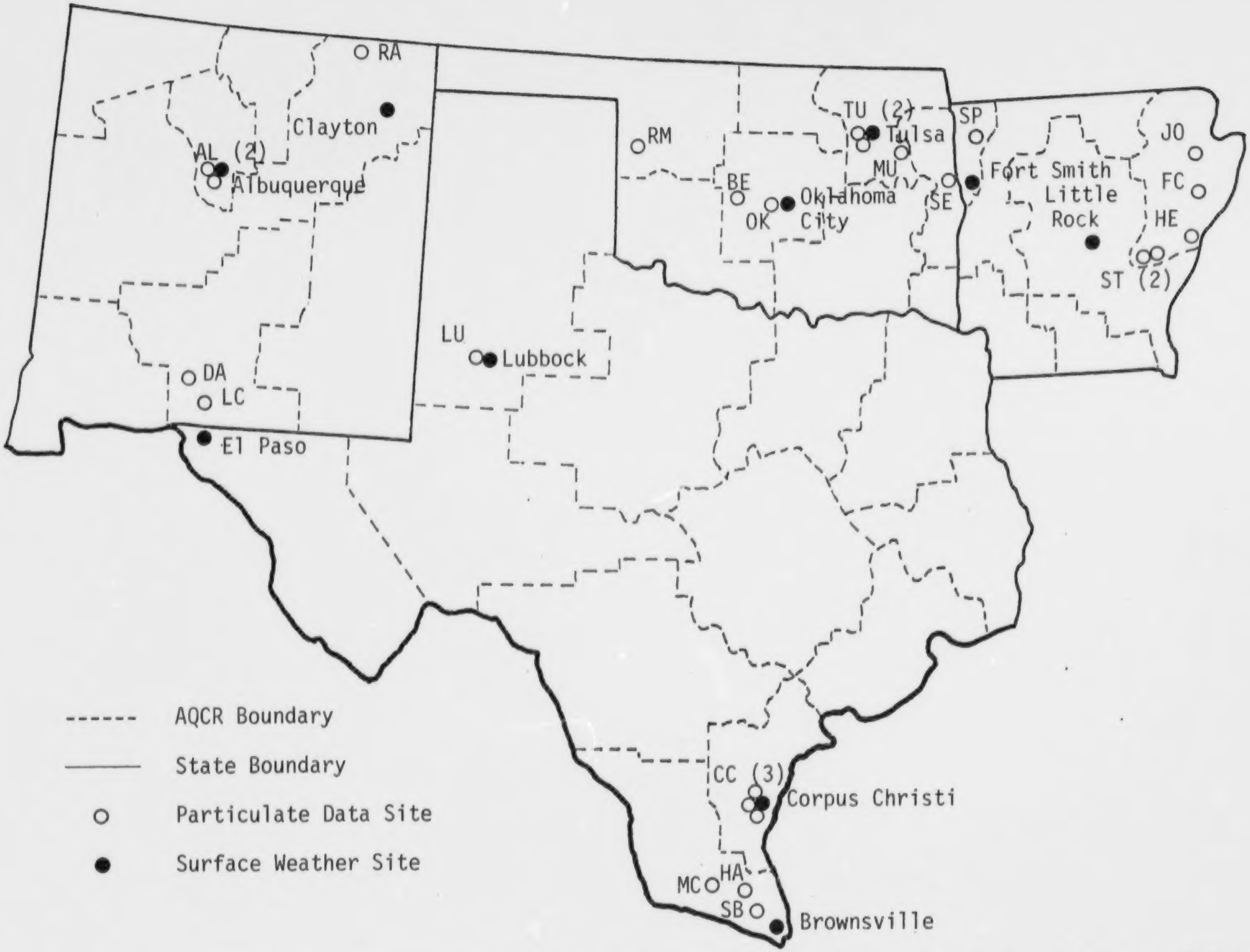

RA
Clayton
AL (2)
Albuquerque
DA
LC
El Paso
LU
Lubbock
RM
BE
OK
Oklahoma City
TU (2)
Tulsa
MU
SE
SP
Fort Smith
Little Rock
JO
FC
HE
ST (2)
CC (3)
Corpus Christi
MC
HA
SB
Brownsville
AQCR Boundary
State Boundary
Particulate Data Site
Surface Weather Site

Figure 2. Location of 25 TSP monitoring stations shown in the regional map of AQCRs (numerals in parenthesis indicate the number of stations in a same city).

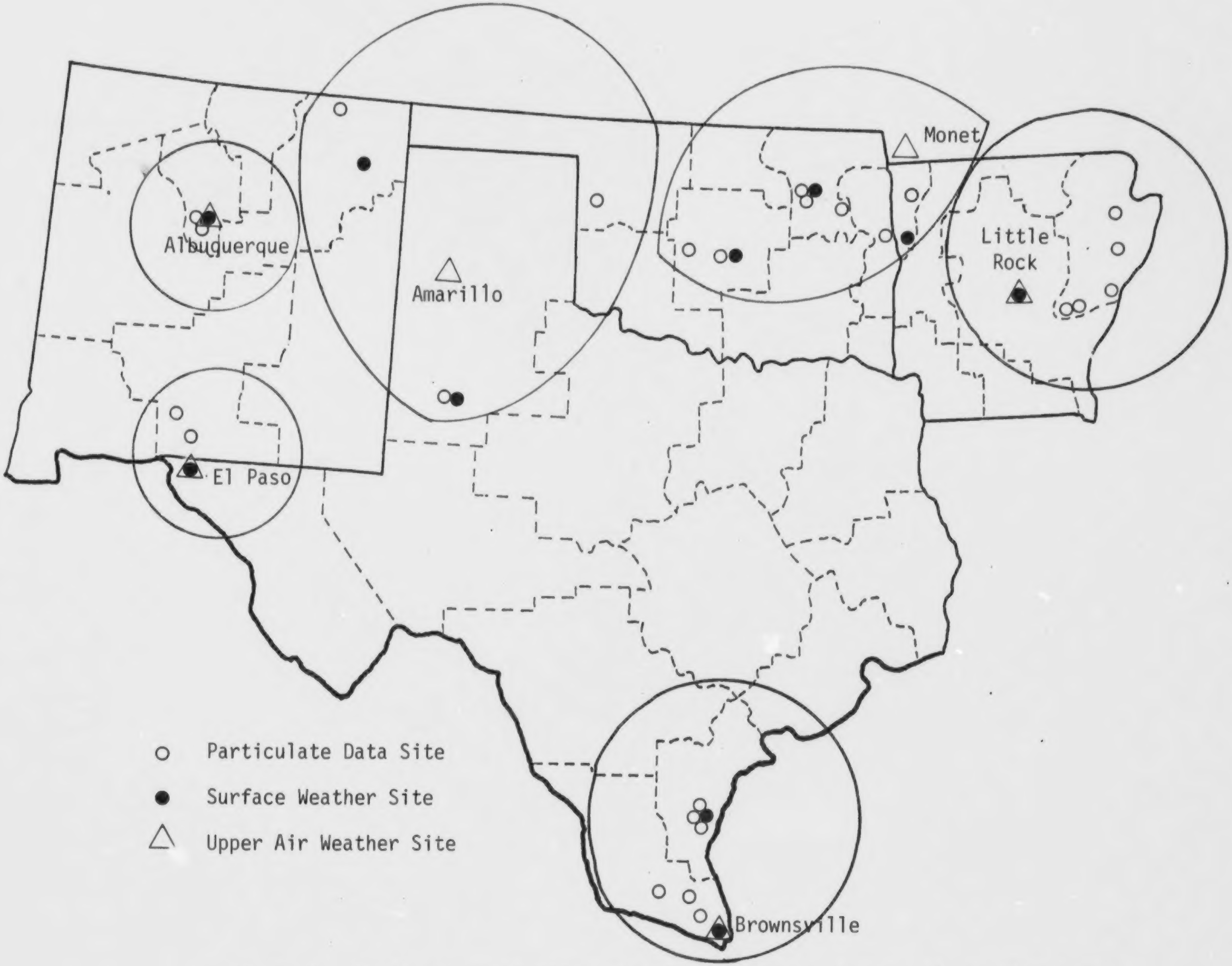

Monet
Albuquerque
Amarillo
Little Rock
El Paso
Brownsville
○ Particulate Data Site
● Surface Weather Site
△ Upper Air Weather Site

Figure 3. Combinations of TSP monitoring stations and upper air weather sites.

TABLE 2. SUMMARY OF SITE CHARACTERISTICS

AQCR	SITE NAME	SAROAD CODE	COUNTY/STATE	SAROAD SITE TYPE	COUNTYWIDE EMISSION DENSITY	CLIMATOLOGY				ELEVATION OF MONITOR ABOVE GROUND
						PRECIPITATION	HUMIDITY	WIND SPEED	HEATING DEGREE DAYS	
211	LUBBOCK (LU)	453340 001F01	Lubbock, TX	S/C	Moderate	M. Low	M. Low	High	Moderate	14 ft.
213	HARLINGEN (HA)	452320 001F01	Cameron, TX	S/C	Moderately Low	Moderate	High	M. High	Low	10
	McALLEN (MC)	453390 001F01	Hidalgo, TX	S/C	Moderately Low	Moderate	High	M. High	Low	30
	SAN BENITO (SB)	454600 001F01	Cameron, TX	CC/C	Moderately Low	Moderate	High	M. High	Low	10
214	CORPUS CHRISTI (CC)	451150 001F01	Nueces, TX	S/R	Moderately High	M. High	High	M. High	Low	20
	CORPUS CHRISTI (CC)	451150 004F02	Nueces, TX	S/I	Moderately High	M. High	High	M. High	Low	14
	CORPUS CHRISTI (CC)	451150 019F01	Nueces, TX	S/R	Moderately High	M. High	High	M. High	Low	13
017	SPRINGDALE (SP)	042480 001F01	Washington, AR	CC/C	Moderate	High	High	M. Low	Moderate	12
	SEQUOYAH (SE)	372760 471F01	Sequoyah, OK	R/I	Moderately Low	M. High	M. High	M. Low	Moderate	10
020	FORREST CITY (FC)	040900 001F01	St. Francis, AR	CC/C	Moderately Low	High	High	M. Low	Moderate	18
	HELENA (HE)	041080 001F01	Phillips, AR	CC/C	Moderately Low	High	High	M. Low	Moderate	10
	JONESBORO (JO)	041320 001F01	Craighead, AR	CC/C	Moderate	High	High	M. Low	Moderate	16
	STUTTGART (ST)	042540 001F01	Arkansas, AR	CC/C	Moderate	High	High	M. Low	Moderate	14
	STUTTGART (ST)	042540 005F05	Arkansas, AR	CC/R	Moderate	High	High	M. Low	Moderate	15
187	ROGER MILLS (RM)	372600 900F01	Roger Mills, OK	R/A	Low	M. Low	Moderate	High	Moderate	12
184	BETHANY (BE)	370260 014F01	Oklahoma, OK	S/R	Moderately High	M. High	M. High	High	Moderate	4
	OKLAHOMA CITY (OK)	372200 002F01	Oklahoma, OK	S/C	Moderately High	M. High	M. High	High	Moderate	70
186	TULSA (TU)	373000 110F01	Tulsa, OK	S/R	Moderately High	M. High	M. High	M. High	Moderate	27
	TULSA (TU)	373000 112F01	Tulsa, OK	CC/C	Moderately High	M. High	M. High	M. High	Moderate	4
	MUSKOGEE (MU)	371980 162F01	Muskogee, OK	S/C	Moderately Low	M. High	M. High	M. High	Moderate	4
152	ALBUQUERQUE (AL)	320040 003H01	Bernalillo, NM	CC/C	High	Low	M. Low	Moderate	M. High	29
	ALBUQUERQUE (AL)	320040 008H02	Bernalillo, NM	S/C	High	Low	M. Low	Moderate	M. High	13
153	DONA ANA (DA)	320340 001F02	Dona Ana, NM	R/C	Low	Low	Low	M. High	Moderate	12
	LAS CURCES (LC)	320580 001F01	Dona Ana, NM	CC/C	Low	Low	Low	M. High	Moderate	50
154	RATON (RA)	320900 001F01	Colfax, NM	S/C	Moderately Low	M. Low	Moderate	Moderate	M. High	15

TABLE 3. SUMMARY OF SITE CLASSIFICATIONS

Site Types	Commercial	Industrial	Residential	Agricultural	Total
Center City	9	0	1	0	10
Suburban	7	1	4	0	12
Rural	1	1	0	1	3
Total	17	2	5	1	25

2.2.2 TSP Emission Density

Particulate emissions arise partly from fugitive dust sources (natural or man-made) and partly from conventional anthropogenic sources such as fuel combustion and industrial processes. Particulate emissions from conventional sources are available on a county-by-county basis from EPA's National Emissions Data System (Pace 1977). Table 4 lists TSP emissions from conventional sources for each of the counties included in our study. The emissions are broken down according to area sources and large point sources; total emission density is also listed for each county.

Bernalillo, NM exhibits an exceptionally high emission density among counties in the study region. The high emission density (57.6 tons/yr./sq.mi.) is attributed to a large amount of TSP emissions from point sources (64,050 tons/yr.). Nueces, TX and Oklahoma, OK and Tulsa, OK have moderately high TSP emission densities, 10 to 20 tons/yr./sq.mi. Lubbock, TX and Washington, Craighead, and Arkansas, AR have moderate emission densities of 5 to 10 tons/yr./sq.mi. Hidalgo, TX, Cameron, TX, Sequoyah, OK, Muskogee, OK, St. Francis, AR, Phillips, AR, and Colfax, NM have moderately low emission densities, ranging from 1 to 5 tons/yr./sq.mi. Two counties, Roger Mills, OK and Dona Ana, NM have low TSP emission densities, less than 1 ton/yr./sq.mi.

2.2.3 Climatic Characteristics

The study region encompasses a large geographical area ranging from humid coastal areas to dry inland plains to mountainous areas.

TABLE 4. ANNUAL TSP EMISSIONS BY COUNTY

County Name	AQCR	Area Sources tons/yr.	Point Sources tons/yr.	Emission Density tons/yr./sq.mi.
Lubbock, TX	211	4,939	959	6.6
Hidalgo, TX	213	1,199	438	1.1
Cameron, TX	213	1,045	1,166	2.5
Nueces, TX	214	10,925	4,728	18.5
Sequoyah, OK	017	455	402	1.2
Roger Mills, OK	187	119	0	.1
Oklahoma, OK	184	5,515	2,406	11.2
Tulsa, OK	186	3,821	2,835	11.6
Muskogee, OK	186	674	462	1.4
Washington, AR	017	785	4,394	5.4
St. Francis, AR	020	417	791	1.9
Phillips, AR	020	356	1,917	3.3
Craighead, AR	020	504	5,101	7.8
Arkansas, AR	020	249	7,820	7.9
Bernalillo, NM	152	3,043	64,050	57.6
Dona Ana, NM	153	763	380	.3
Colfax, NM	154	187	10,308	2.7

Figure 4 shows the boundaries of state climatic divisions and the locations of the 25 monitoring sites. This section uses data from the National Climatic Atlas (ESSA 1968) to identify the climate characteristics of each monitoring location.

Precipitation

Annual total precipitation in each state climatic division is illustrated in Figure 5. The six Arkansas sites have annual precipitation of 45 inches or more (high). The Oklahoma sites (except for Roger Mills) and the three Corpus Christi sites have annual precipitation between 30 and 45 inches (moderately high). Harlingen, McAllen, and San Benito sites in AQCR 213 experience moderate precipitation, 20 to 30 inches. Roger Mills, OK, Lubbock, TX and Raton, NM have annual precipitation between 15 and 20 inches (moderately low). The four New Mexico sites at Albuquerque, Dona Ana, and Las Cruces have low annual precipitation, less than 15 inches.

In addition to total precipitation, frequency of measurable precipitation may be important to characterize the dryness of a monitoring site. Figure 6 shows the annual mean number of days with 0.01 inch or more of precipitation. All the Arkansas sites have more than 100 such days, while the Oklahoma sites except for Roger Mills have between 80 and 100 days. Roger Mills and the stations in Texas and New Mexico have less than 80 such days.

Relative Humidity

Figure 7 shows an isopleth map of annual average relative humidity. All the sites in Arkansas and Texas except for Lubbock have annual relative humidity exceeding 70% (high). The Oklahoma sites except Roger Mills exhibit moderately high relative humidity, 60 to 70%. Roger Mills and Raton have annual relative humidity just below 60% (moderate). Lubbock and the two Albuquerque locations have moderately low relative humidity, just above 50%. Dona Ana and Las Cruces have low relative humidity, about 40%.

Wind Speed

Annual surface wind roses at several locations in the study region are presented in Figure 8. Southeasterly winds are dominant in Oklahoma

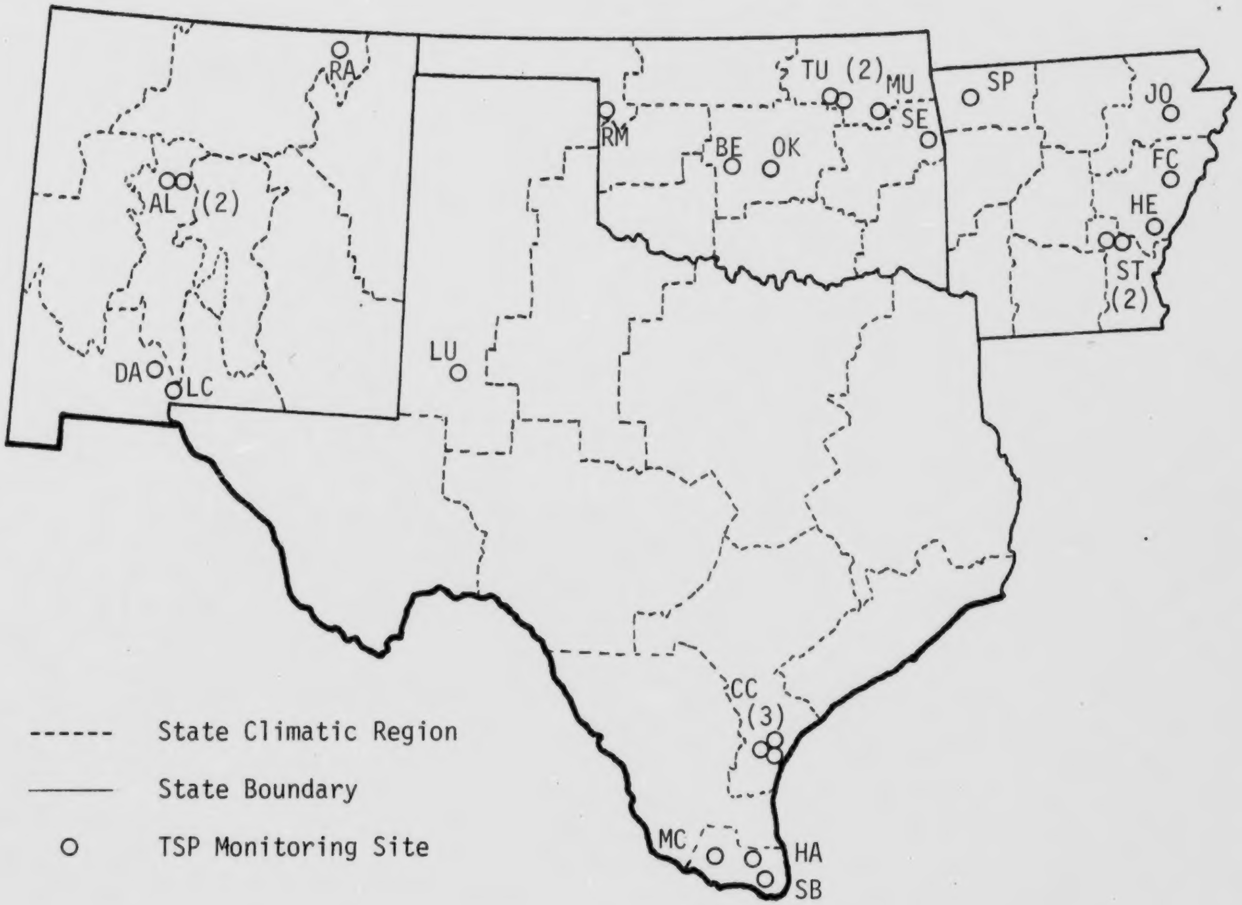

RA
AL (2)
DA
LC
LU
RM
BE
OK
TU (2)
MU
SE
SP
JO
FC
HE
ST (2)
CC (3)
MC
HA
SB
State Climatic Region
State Boundary
TSP Monitoring Site

Figure 4. Location of 25 TSP monitoring stations shown in the Regional Map of State Climatic Divisions (numerals in parenthesis indicate the number of stations in a same city).

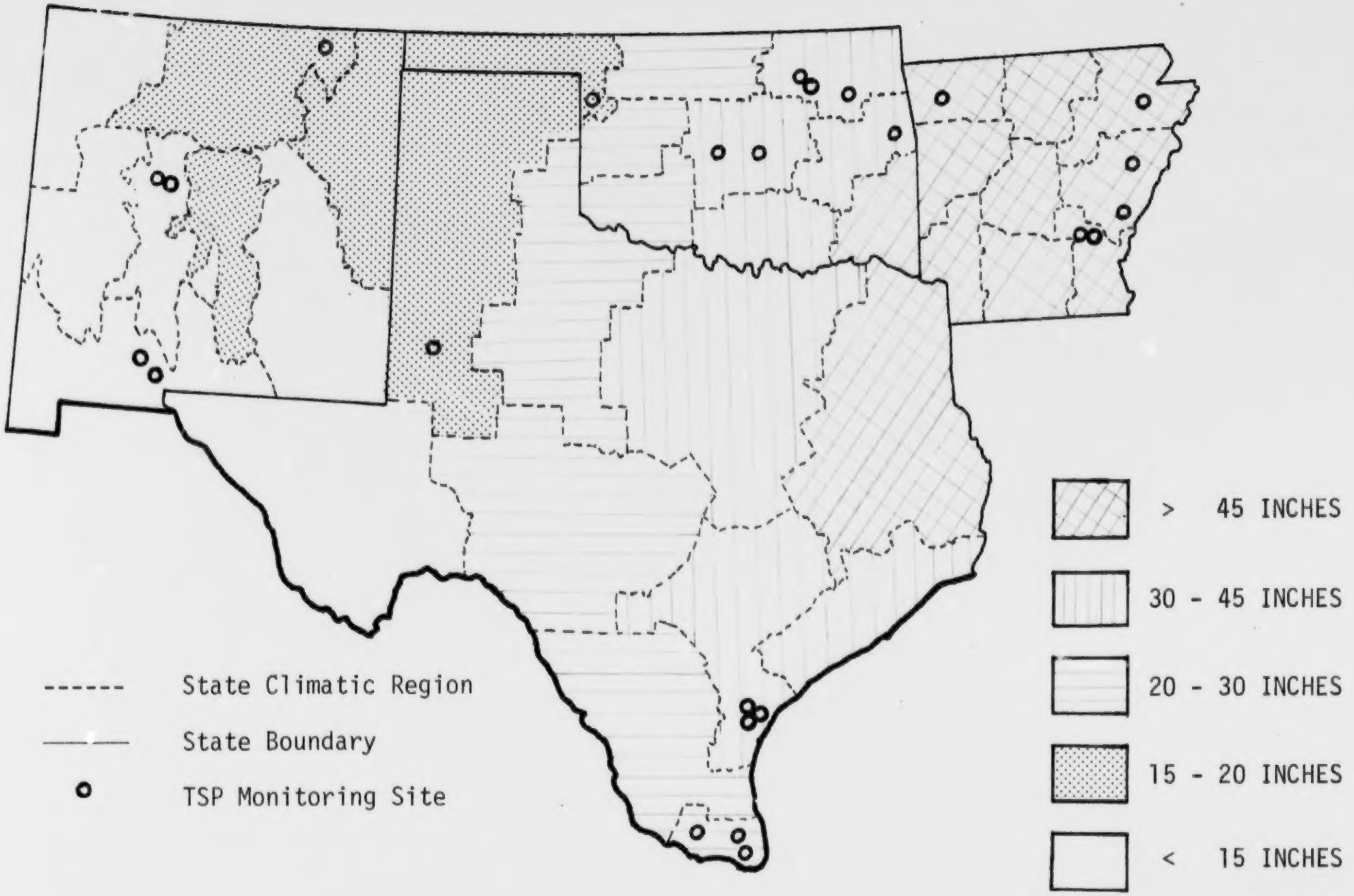

State Climatic Region
State Boundary
TSP Monitoring Site
> 45 INCHES
30 - 45 INCHES
20 - 30 INCHES
15 - 20 INCHES
< 15 INCHES

Figure 5. Mean annual precipitation, based on the period 1931 - 55.

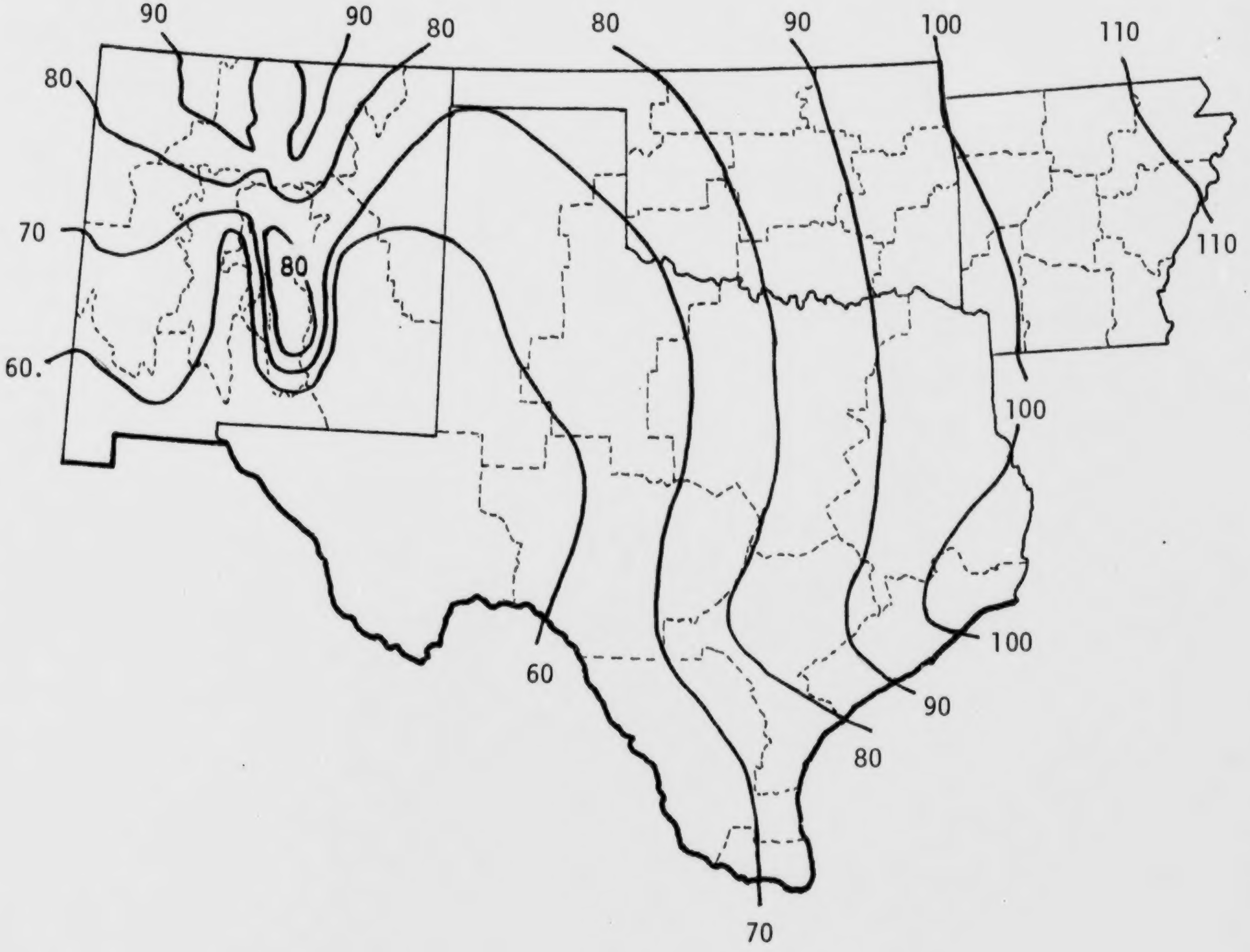

90
90
80
80
90
100
110
80
70
80
60.
110
100
100
60
90
80
70

Figure 6. Annual number of days with 0.01 inch or more of precipitation.

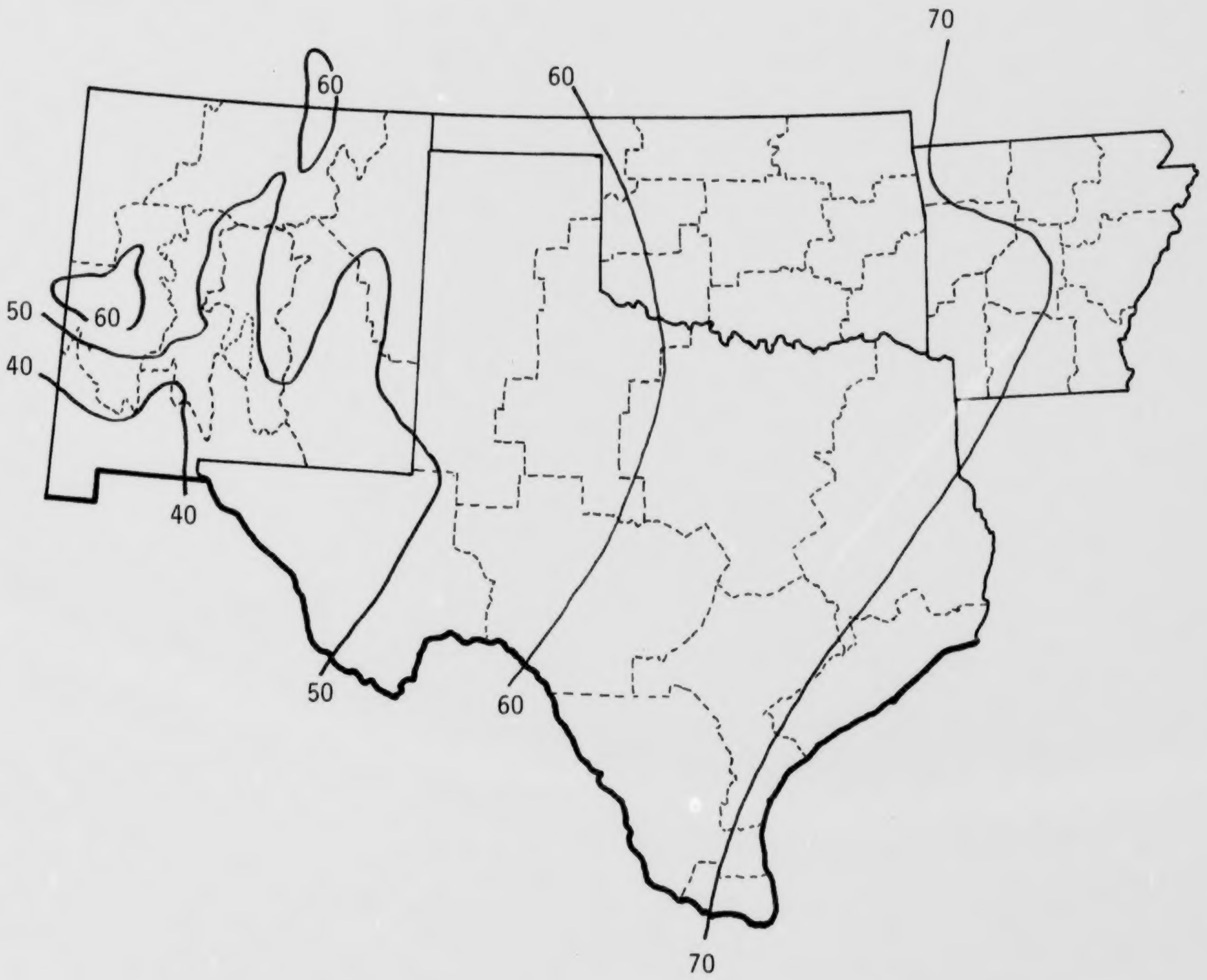

70
60
60
50
60
40
40
50
60
70

Figure 7. Mean annual relative humidity in percent.

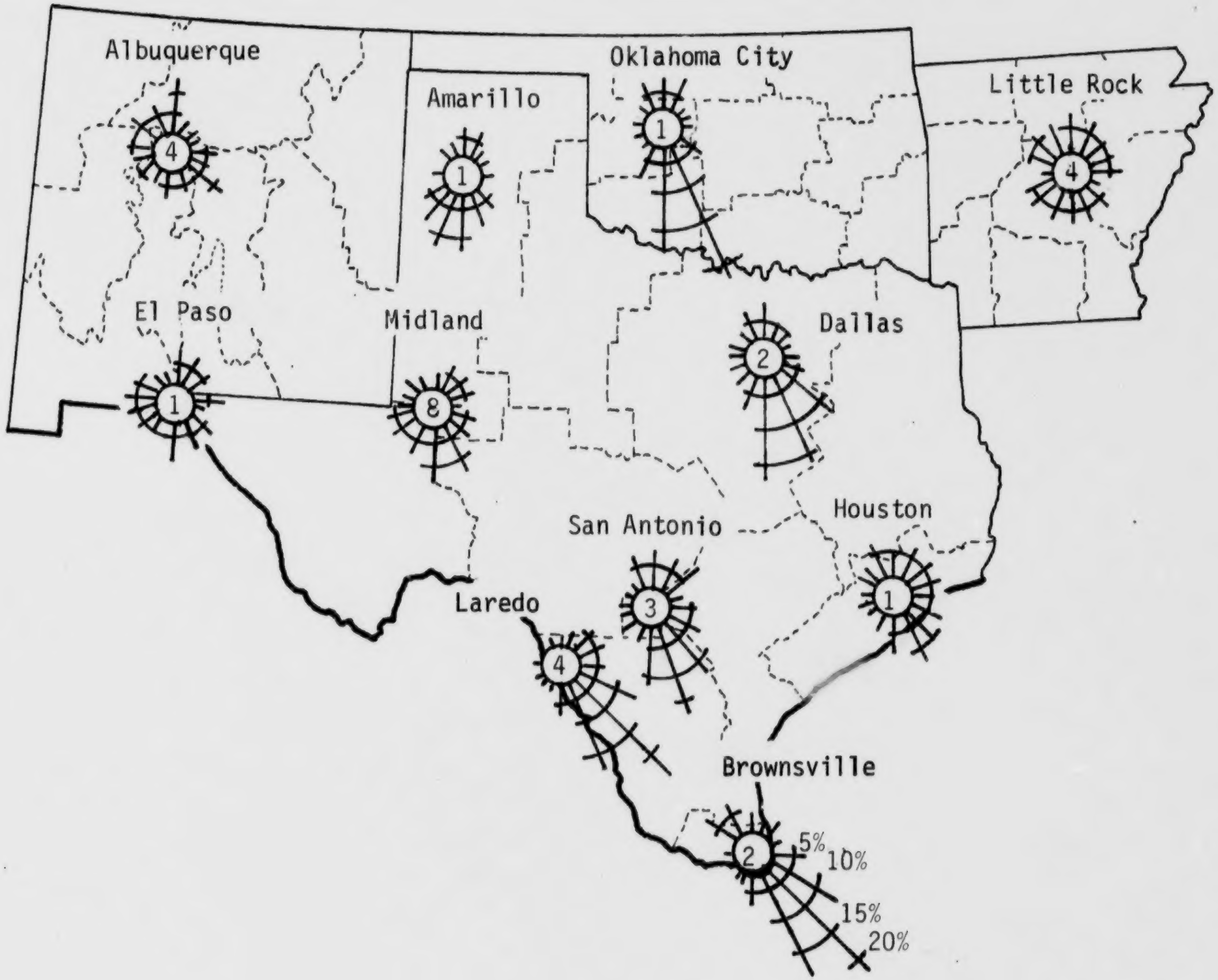

Albuquerque
4
Amarillo
1
Oklahoma City
1
Little Rock
4
El Paso
1
Midland
8
Dallas
2
San Antonio
3
Houston
1
Laredo
4
Brownsville
2
5%
10%
15%
20%

Figure 8. Annual surface wind roses.

and most of Texas. Wind direction is uniformly distributed in Oklahoma, New Mexico, and western Texas.

Figure 9 illustrates prevailing wind directions and annual average wind speeds. Average wind speeds are high (about 13 to 14 m.p.h.) at Lubbock, Roger Mills, Bethany, and Oklahoma City. Average wind speeds are moderately high (11 to 12 m.p.h.) at the six sites in Southern Texas, the two sites in southern New Mexico, and the three sites near Tulsa, Oklahoma. Moderate wind speeds (9 to 10 m.p.h.) occur at the three sites in northern New Mexico, while the six Arkansas sites and Sequoyah, OK experience moderately low wind speeds (about 8 m.p.h.).

Heating Degree Days

Figure 10 depicts isopleths of annual total heating degree days. The three sites in northern New Mexico have moderately high values, between 5000 and 6000 degree days. The six sites in southern Texas have low values, less than 1000 degree days. All the other sites have moderate values, between 3000 and 4000 degree days.

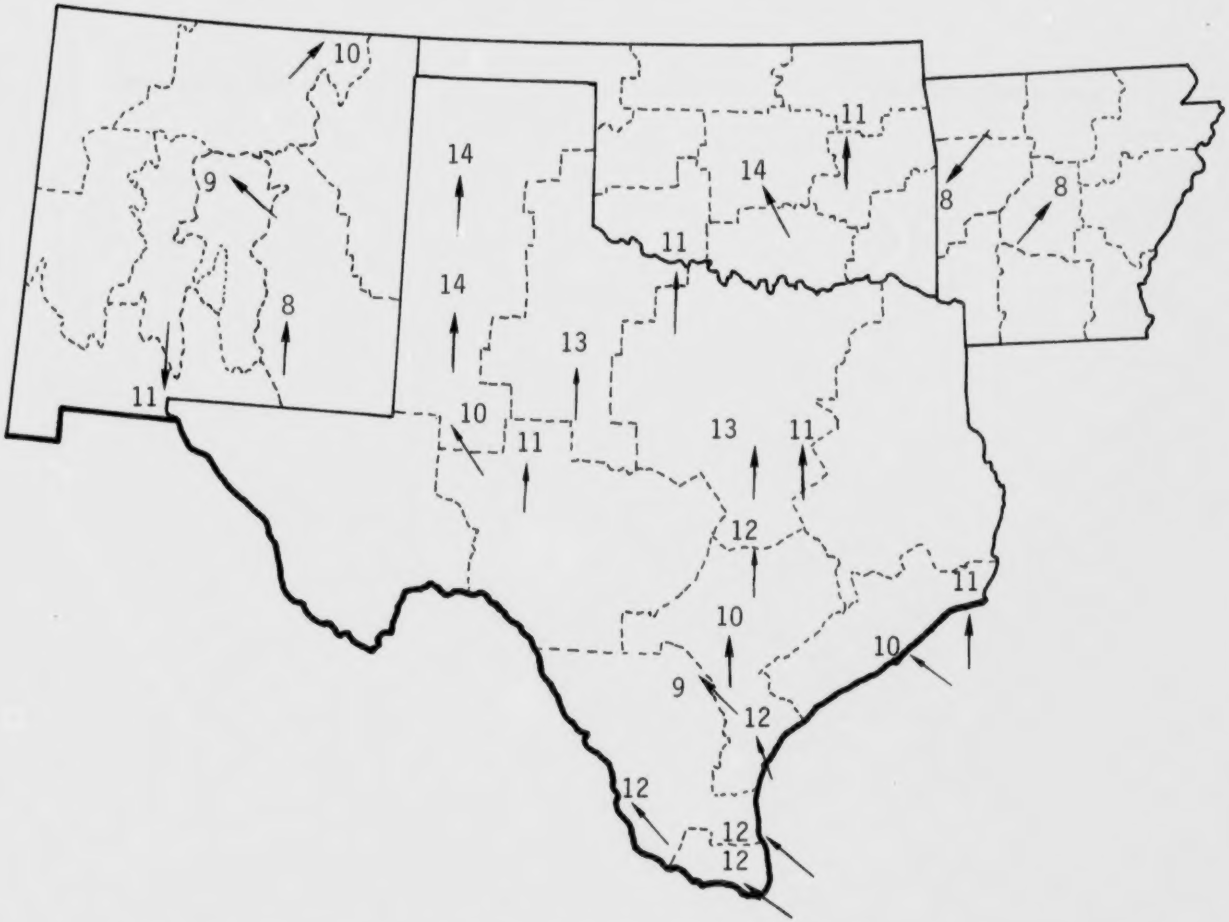

10
9
8
11
14
14
10
11
13
11
14
11
13
11
12
10
9
12
11
10
12
12
12
8
8

Figure 9. Prevailing surface wind direction and annual mean wind speed (m.p.h.).

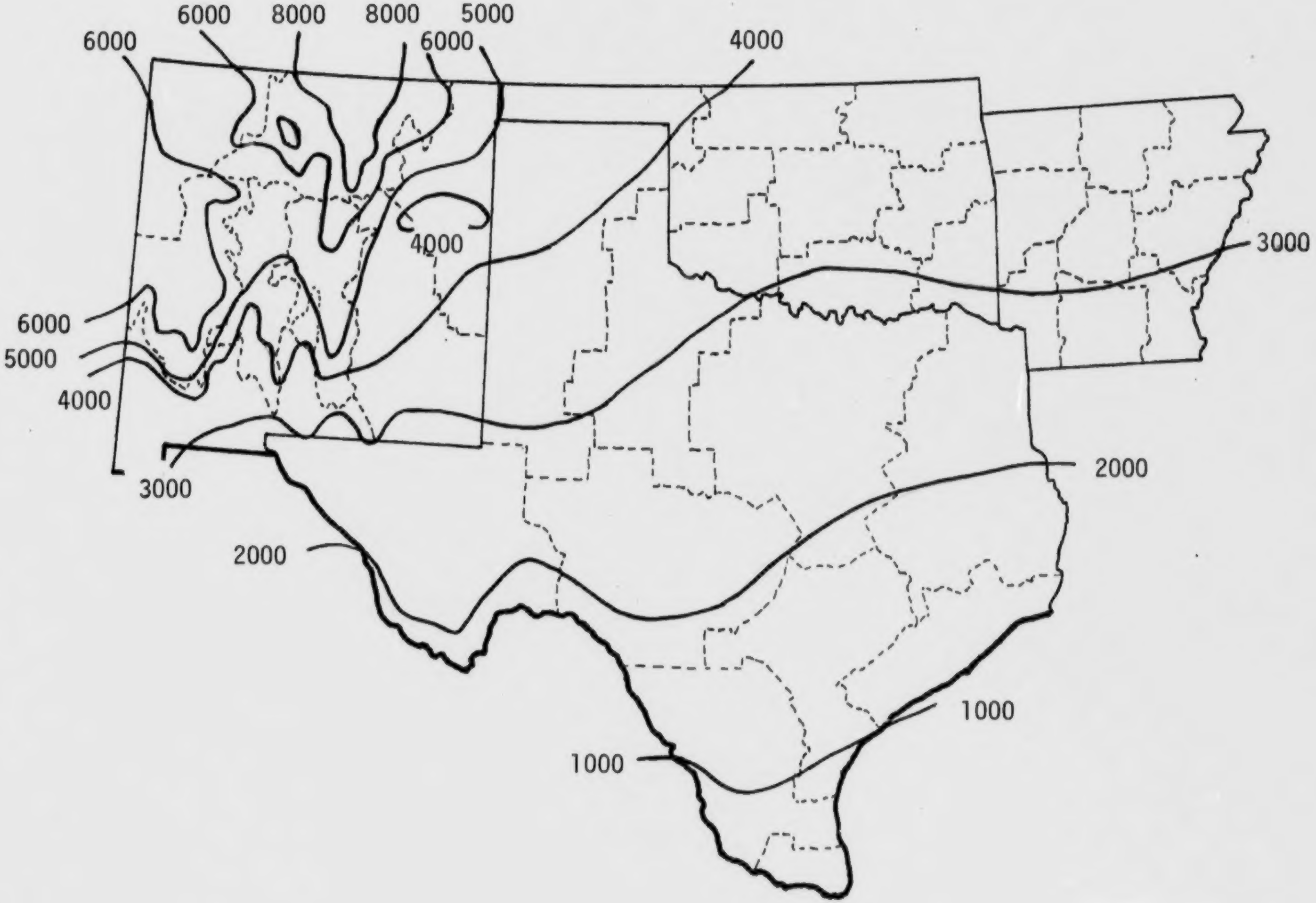

6000
6000
8000
8000
5000
6000
4000
4000
3000
6000
5000
4000
3000
2000
2000
1000
1000

Figure 10. Annual total heating degree days.

3.0 QUALITY CHECK OF TSP DATA

In order to help eliminate erroneous TSP values, a quality check was conducted on the TSP data from all 25 sites. Outliers in the TSP data were identified by three methods: (1) by selecting all values outside upper (500 $\mu g/m^3$) and lower (10 $\mu g/m^3$) thresholds; (2) by examining the upper tail of the cumulative TSP frequency distribution for each site; and (3) by comparing TSP data values at different sampling sites for each sampling date. All outliers detected by these three methods were checked by calling the relevant local monitoring agency and by examining meteorological conditions on the days when the outliers occurred. Detailed discussions of the detection methods and confirmation procedures follow.

3.1 OUTLIER DETECTION USING UPPER AND LOWER THRESHOLDS

A cursory scan of TSP data summaries (EPA 1974) indicated that the great majority of 24-hour TSP values occur within the range 10 to 500 $\mu g/m^3$. All values in our data base that fall outside this range were printed out. Table 5 lists the outliers detected by the upper and lower thresholds.

We suspected that the zeroes in Table 5 might represent missing data rather than zero TSP concentrations; later analysis confirmed these suspicions. The other low values in Table 5 are further examined in the section on site-to-site comparison of TSP data. The high TSP values are examined in the sections on percentile concentration plots and site-to-site comparisons.

3.2 OUTLIER DETECTION USING PERCENTILE CONCENTRATION PLOTS

The TSP data for each monitoring site were sorted in ascending order of magnitude to obtain a cumulative frequency distribution. The rank of each data value was then divided by N + 1 (N = sample size) to obtain a percentile concentration distribution. The percentile concentrations were then analyzed using the lognormal distribution and the exponential distribution.

The mathematical function most often used to analyze air pollution frequency distributions is the lognormal distribution popularized by Larsen and his co-workers (Zimmer and Larsen 1965; Larsen et al 1967; Larsen 1969,

TABLE 5. OUTLIERS OF TSP DATA IN REGION 6 (ALL VALUES IN $\mu g/m^3$)

State	City	Site Code	Date of Occurrence	Data Value
Arkansas	Helena	041080 001F01	75-11-02	9
"	Stuttgart	042540 005F05	73-08-26	512
"	"	" "	73-10-19	924
New Mexico	Albuquerque	320040 003H01	75-08-10	5
"	"	320040 008H02	73-01-04	7
"	"	" "	73-01-16	6
"	"	" "	73-01-28	0
"	"	" "	74-03-12	3
"	Dona Ana	320340 001F02	74-09-18	7
"	Las Cruces	320580 001F01	75-11-14	503
Oklahoma	Muskogee	371980 162F01	74-01-11	5
"	"	" "	74-01-17	9
"	Oklahoma City	372200 002F01	73-02-01	4
"	"	" "	74-03-10	8
"	Sequoyah	372760 471F01	73-01-01	7
"	"	" "	73-01-19	8
"	"	" "	73-01-23	7
"	"	" "	73-03-05	9
"	Tulsa	373000 112F01	73-01-28	0
"	"	" "	73-02-09	0
"	"	" "	73-03-07	0
"	"	" "	73-03-29	0
"	"	" "	73-04-22	0
"	"	" "	73-04-26	0
"	"	" "	73-06-19	0
"	"	" "	73-07-02	0
"	"	" "	73-08-16	0
"	"	" "	73-09-13	0
"	"	" "	75-01-18	561
Texas	San Benito	454600 001F01	73-07-15	1302
"	"	" "	75-07-11	883
"	"	" "	75-07-29	791
"	"	" "	75-09-03	549

1971, 1973, 1974). When properly used, the lognormal distribution can be a valuable tool in analyzing data for TSP and other pollutants (Gifford 1974; Singpurwalla 1972; Neustadter et al. 1973; Trijonis 1977; Richard et al. 1977). If TSP data are lognormally distributed, the percentile concentration plot on log-normal probability paper should resemble a straight line. An example of a good lognormal fit is shown in Figure 11, representing TSP data at Jonesboro, Arkansas.

Because plots on semi-lograthmic paper emphasize the upper tail of the distribution, the exponential distribution function is gaining popularity in the analysis of air quality data (Curran and Frank 1975; Meisel and Dushane 1977). The cumulative exponential distribution function is given by:

$$F(x) = 1 - \exp(-\lambda x) \qquad x > 0. \qquad (1)$$

If TSP data are exponentially distributed, the percentile concentration plot should appear as a straight line on semi-log paper. Figure 12 presents an example of such a plot and indicates that the TSP data at Corpus Christi (004F02) closely follow an exponential distribution.

Lognormal and exponential plots were prepared for each of the 25 TSP sites, using combined data for 1973-1975 (see Appendix A). Values that deviated significantly from the general shapes of the distributions were considered as outliers. Figures 13 and 14 illustrate the detection of outliers for Stuttgart (001F01) using lognormal and exponential plots, respectively. All outliers detected using the cumulative frequency plots were subjected to the confirmation precedures discussed in Section 3.4.

3.3 OUTLIER DETECTION FROM SITE-BY-SITE COMPARISONS

As shown in Table 1, there are several groups of TSP sites that are associated with a single nearby weather station. For each of these groups, all TSP data values were printed out for each day on which at least one station reported a TSP measurement. These printouts facilitated an examination of the data for consistency among the stations. When one station deviated significantly from the others, or when any station reported extremely high or low values (as in Table 5), the data were selected for the confirmation processes.

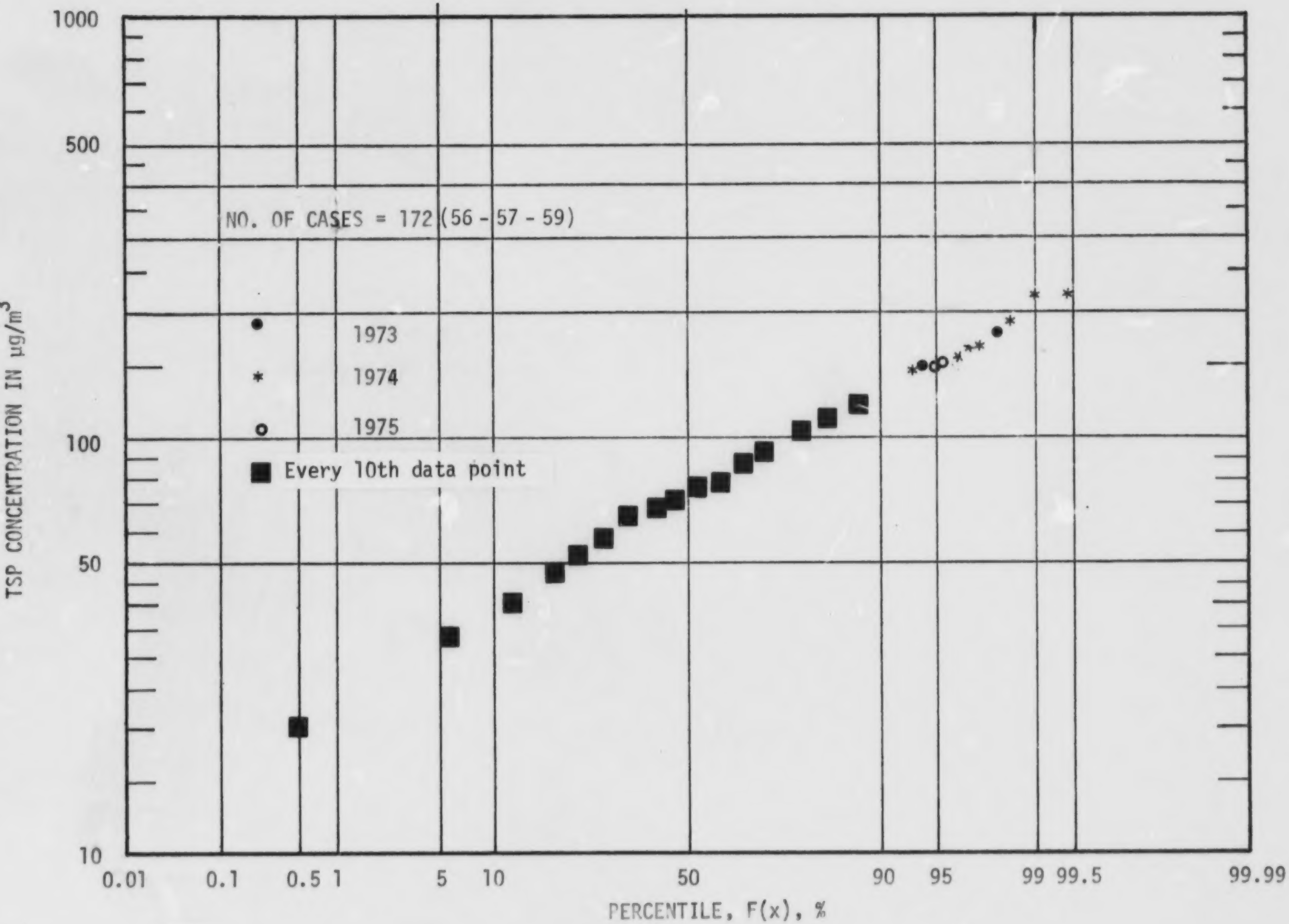


Figure 11. Log-normal distribution plot of TSP data at Jonesboro, Arkansas (#041320001F01).

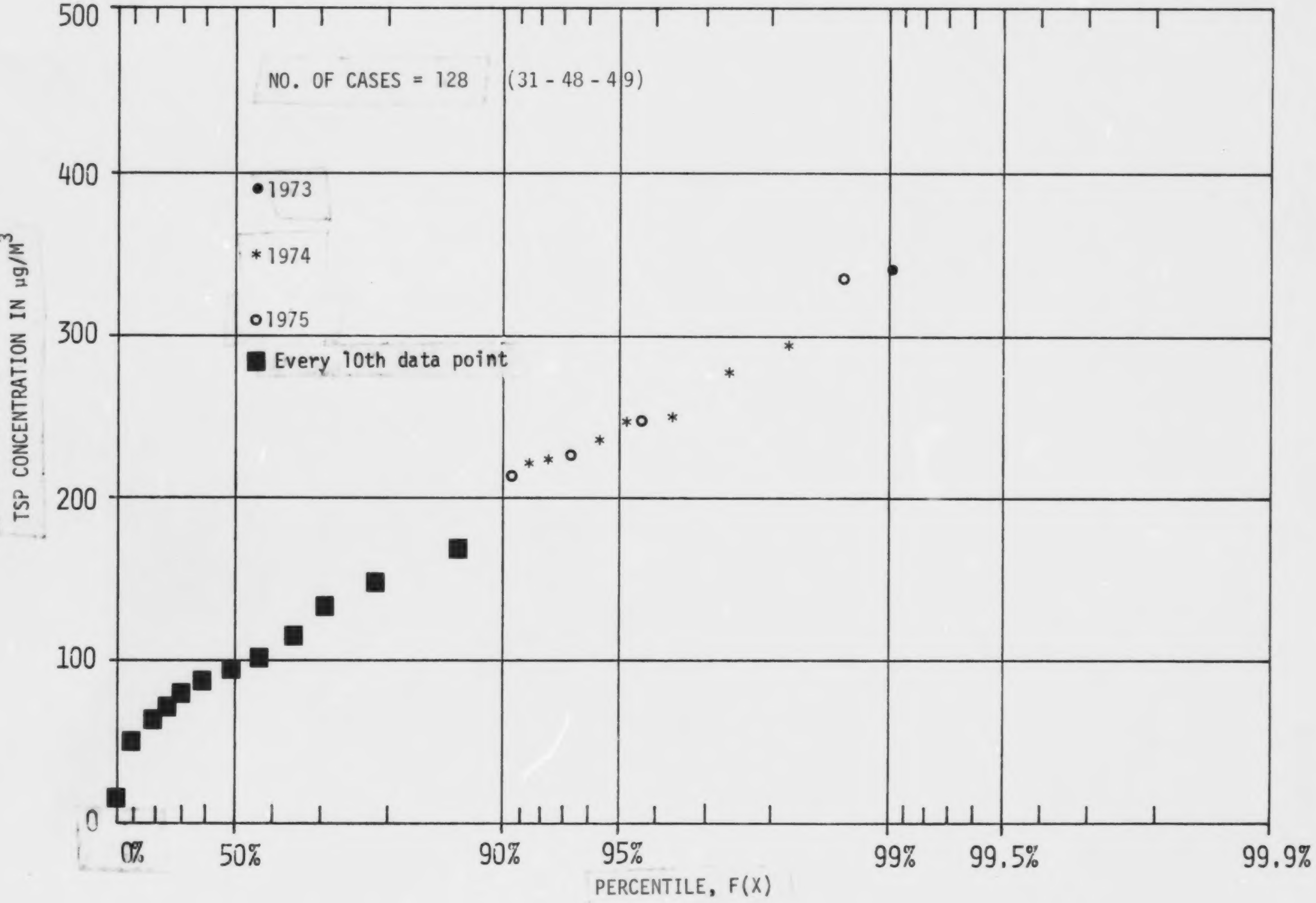


Figure 12. Exponential distribution plot of TSP data at Corpus Christi, TX (#451150004F02)

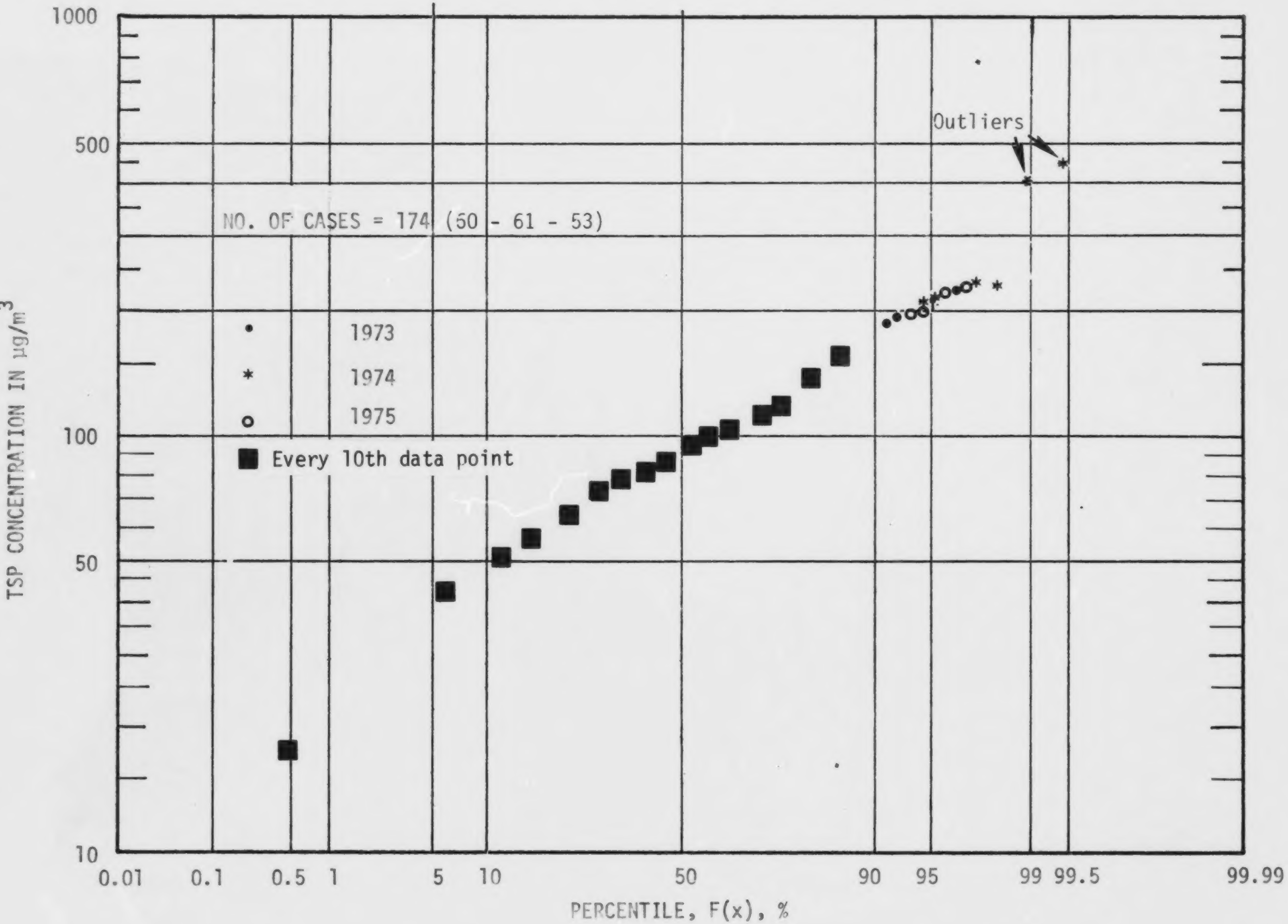


Figure 13. Log-normal distribution plot of TSP data at Stuttgart, Arkansas (#042540001F01).

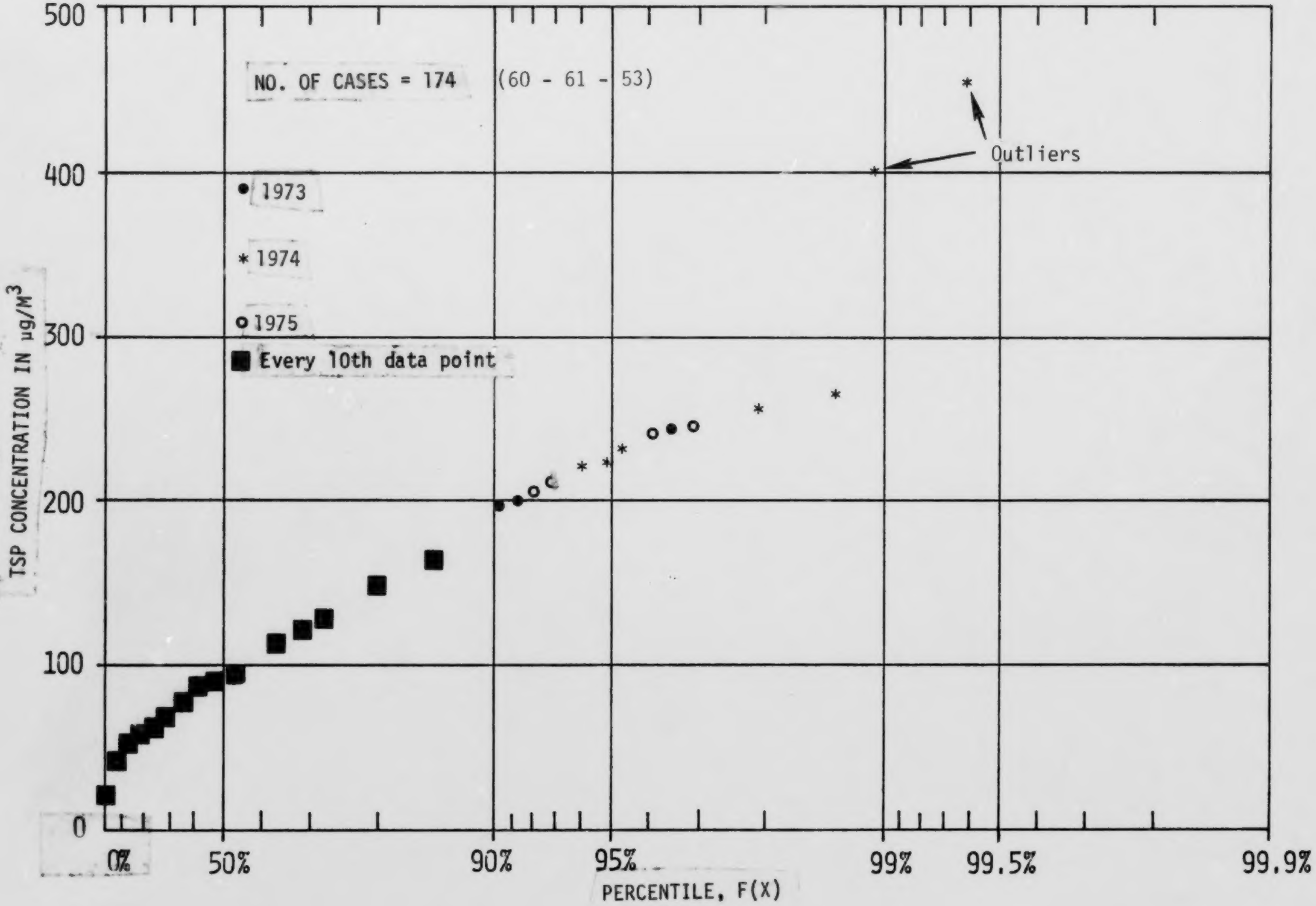


Figure 14. Exponential distribution plot of TSP data at Stuttgart, AR (#042540001F01)

Table 6 lists outliers detected by the visual scan of TSP data. Table 6 also indicates the meteorological features on the days when the outliers occurred. The meteorological features were obtained from the "Local Climatological Data" summaries for the associated surface weather stations.

As evident in Table 6, the records for Tulsa site 112F01 include several zero readings in 1973. In one sense these readings are plausible; the zero readings usually occur on days after heavy storms or on days with trace precipitation (the same meteorology associated with several low, non-zero readings). However, it is also possible that the heavy storms on the previous days interrupted the sampling plan, causing a scheduled measurement to be cancelled. Conversations with the Oklahoma Air Quality Service indicate the latter is the case, i.e. the zeroes represent missing data.

Several of the high TSP readings are verified to some extent by reports of either smoke/haze or high winds. Also, when one station (e.g. Corpus Christi 004F02) reports an exceptionally high datum compared to nearby sites, we usually find that the station generally records higher readings than nearby sites (making the exceptionally high datum more plausible).

3.4 CONFIRMATION OF OUTLIERS

All the TSP data points flagged by the three different outlier detection methods were subjected to confirmation procedures. We called the local monitoring agencies and inquired about the validity of the flagged values. The agency personnel that were contacted are listed in Table 7.

Except for the zero readings, the exceptionally low TSP values were verified. The zero SAROAD value at Albuquerque (008H02) on 1-28-73 was actually 11 $\mu g/m^3$, evidently a transcription error. The ten zero readings at Tulsa (112F01) in 1973 were all missing values.

The outliers involving extremely high TSP levels are listed in Table 7. As indicated by Table 7, the great majority of the outliers were confirmed by the telephone survey. These outliers also appeared to be plausible based on meteorological conditions and on site-to-site comparisons of TSP data. There are two values which were invalidated by the telephone survey. The recording of 503 $\mu g/m^3$ on 11-14-75 at the Las Cruces station has been changed to 228 $\mu g/m^3$; the recording of 561 $\mu g/m^3$ on 1-18-75 at Tulsa (112F01) has been changed to 186 $\mu g/m^3$.

TABLE 6. OUTLIERS DETECTED BY SITE-BY-SITE COMPARISONS OF TSP DATA

Station #	Station	TSP ($\mu g/m^3$)	Date	Weather
373000112F01 371980162F01	Tulsa 2 Muskogee	0 22	1/28/73	Cold - with a trace of snow
373000112F01 371980162F01	Tulsa 2 Muskogee	0 66	2/09/73	Coldest min-temp with below normal winds
373000110F01 373000112F01	Tulsa 1 Tulsa 2	66 0	3/07/73	Fog - light wind the day after heavy rains
373000110F01 373000112F01	Tulsa 1 Tulsa 2	62 0	4/16/73	Avg. winds, moderate temp, trace of precipitation
373000110F01 373000112F01	Tulsa 1 Tulsa 2	46 0	4/26/73	Moderate conditions with a trace of precipitation
373000112F01	Tulsa 2	0	4/22/73	Rain .14" and the day after heavy thunderstorm
373000112F01	Tulsa 2	0	3/29/73	Trace precipitation and day after thunderstorm
373000112F01	Tulsa 2	0	6/19/73	Thunderstorm - rain 0.71"
373000112F01	Tulsa 2	0	7/02/73	Avg. day: clear sky,T - 90 w - 12 mph
373000110F01 373000112F01	Tulsa 1 Tulsa 2	54 0	8/16/73	Day after a thunderstorm (otherwise average)
373000112F01	Tulsa 2	0	9/13/73	Fog, light winds and the day after a thunderstorm
371980162F01	Muskogee	5	1/11/74	Fog and a trace of snow
373000110F01 373000112F01 371980162F01	Tulsa 1 Tulsa 2 Muskogee	72 561 46	1/18/75	Heavy fog and hot
371980162F01	Muskogee	9	1/17/74	Trace of precipitation, fog
372760471F01	Sequoyah	7	1/01/73	Avg. temp, light wind, clear skies

TABLE 6. OUTLIERS DETECTED BY SITE-BY-SITE COMPARISONS OF TSP DATA (Con't)

Station #	Station	TSP ($\mu g/m^3$)	Date	Weather
372760471F01	Sequoyah	8	1/19/73	Warm, light wind, day after thunderstorm
372760471F01	Sequoyah	7	1/23/73	Avg. conditions but 1.4" rain over 5 previous days
372760471F01	Sequoyah	9	3/05/73	Overcast, avg. temp, avg. wind, day after 4 day rainstorm
042480001F01	Springdale	31		
040900001F01	Forrest City	374	2/21/73	Moderate temp and average winds
041080001F01	Helena	99		
041320001F01	Jonesboro	55		
042540001F01	Stuttgart 1	149		
040900001F01	Forrest City	103	8/26/73	Hot, smoke, haze, average winds
041080001F01	Helena	88		
041320001F01	Jonesboro	104		
042540001F01	Stuttgart 1	99		
042540005F05	Stuttgart 2	512		
040900001F01	Forrest City	161	10/19/73	Smoke and haze observed along with light winds and avg. temp (harvest time)
041080001F01	Helena	102		
041320001F01	Jonesboro	146		
042540001F01	Stuttgart 1	187		
042540005F05	Stuttgart 2	924		
040900001F01	Forrest City	52	11/02/75	Overcast, trace precipitation, light wind, high humidity
041080001F01	Helena	9		
041320001F01	Jonesboro	21		
042540001F01	Stuttgart 1	58		
042540005F05	Stuttgart 2	161		
451150001F01	Corpus Christi 1	76	2/04/74	Cool, windy and clear
451150004F02	Corpus Christi 2	213		
451150019F01	Corpus Christi 3	47		
451150001F01	Corpus Christi 1	88	3/18/74	Fog and trace precip. avg. wind and overcast
451150004F02	Corpus Christi 2	248		
451150019F01	Corpus Christi 3	43		
451150001F01	Corpus Christi 1	85	6/28/74	Warm, light winds and clear
451150004F02	Corpus Christi 2	188		
451150019F01	Corpus Christi 3	79		

TABLE 6. OUTLIERS DETECTED BY SITE-BY-SITE COMPARISONS OF TSP DATA (Con't)

Station #	Station	TSP (μg/m^3)	Date	Weather
451150001F01	Corpus Christi 1	106	12/08/75	Warm, fog, smoke and haze with light winds
45110004F02	Corpus Christi 2	332		
451150019F01	Corpus Christi 3	74		
454600001F01	San Benito	290	6/28/73	Warm, fog, light winds with smoke and haze on the previous day
452320001F01	Harlingen	66		
453390001F01	McAllen	320		
454600001F01	San Benito	170	7/03/73	Rain and light winds
452320001F01	Harlingen	40		
453390001F01	McAllen	72		
454600001F01	San Benito	424	7/09/73	Warm, fog, avg. winds, smoke and haze
452320001F01	Harlingen	126		
454600001F01	San Benito	1302	7/15/73	Hot - moderate sky and avg. winds
452320001F01	Harlingen	102		
454600001F01	San Benito	421	6/17/75	High winds, clear sky and avg. temp
452320001F01	Harlingen	138		
453390001F01	McAllen	172		
454600001F01	San Benito	240	6/29/75	Trace precip., avg. winds and temp
452320001F01	Harlingen	106		
453390001F01	McAllen	78		
454600001F01	San Benito	316	7/05/75	Light winds and clear skies
452320001F01	Harlingen	85		
454600001F01	San Benito	883	7/11/75	Hot, humid, avg. winds and slight overcast
452320001F01	Harlingen	148		
453390001F01	McAllen	116		
454600001F01	San Benito	791	7/29/75	Hot and humid with average winds
452320001F01	Harlingen	127		
453390001F01	McAllen	112		
454600001F01	San Benito	549	9/03/75	Fog, light winds, hot and humid
453390001F01	McAllen	113		
453340001F01	Lubbock	175	7/22/74	Hot-clear, light winds Trace precipitation
320900001F01	Raton	18		

TABLE 6. OUTLIERS DETECTED BY SITE-BY-SITE COMPARISONS OF TSP DATA (Con't)

Station #	Station	TSP ($\mu g/m^3$)	Date	Weather
453340001F01	Lubbock	108	7/28/74	Hot clear light winds
320900001F01	Raton	32		Trace precipitation
453340001F01	Lubbock	140	8/03/74	Trace precip. cool and above average winds
320900001F01	Raton	42		Heavy rain
372600900F01	Roger Mills	48		Warm, trace precip., light winds
453340001F01	Lubbock	158	8/15/74	Hot, high winds, clear
320900001F01	Raton	65		Trace precipitation
372600900F01	Roger Mills	347		Above avg. winds, hot
320580001F01	Las Cruces	503	11/14/75	Smoke and haze
320040003H01	Albuquerque 1	5	8/10/75	Thunderstorm in the area with a trace of precip.
320040008H02	Albuquerque 2	55		
320040008H02	Albuquerque 2	7	1/04/73	Light winds, avg. temp. - no precipitation
320040008H02	Albuquerque 2	6	1/16/73	High temperature, light winds, clear
320040008H02	Albuquerque 2	0	1/28/73	Light winds, cold night with accumulation of 1" snow pellets on the ground
320040003H01	Albuquerque 1	84	3/12/74	Average temperature, average winds
320040008H02	Albuquerque 2	3		
320040003H01	Albuquerque 1	38	6/10/74	Slight wind, high temp., and low humidity
320040008H02	Albuquerque 2	14		
320340001F01	Dona Ana Co.	27	11/26/74	Cold, cloudy, below average winds
320580001F01	Las Cruces	340		
320340001F01	Dona Ana Co.	255	12/08/75	Smoke, haze, and light winds
320580001F01	Las Cruces	107		
320340001F02	Dona Ana Co.	7	9/18/74	Raining (.24"), warm
320580001F01	Las Cruces	26		

TABLE 7. CHECK OF TSP DATA FLAGGED AS OUTLIERS BY TELEPHONE INQUIRY (ALL TSP VALUES IN $\mu g/m^3$)

Site Name	SAROAD Code	Sampling Date	TSP Value	Remarks
TEXAS	Mr. Larry Butts (512) 451 - 5711 Extension 301 Texas Air Control Board			
Corpus Christi	451150 001F01	4/18/75	265	verified
Corpus Christi	451150 004F02	3/11/75	332	verified
"	451150 004F02	12/08/75	334	verified
Corpus Christi	451150 019F01	4/18/75	201	verified
Harlingen	452320 001F01	8/21/74	298	verified
Lubbock	453340 001F01	12/12/73	339	verified
"	453340 001F01	4/11/74	340	verified
"	453340 001F01	7/10/74	289	verified
"	453340 001F01	4/18/75	496	verified
McAllen	453390 001F01	6/28/73	320	may be invalid
"	453390 001F01	2/22/74	291	verified
San Benito	454600 001F01	7/15/73	1302	verified
"	454600 001F01	7/11/75	883	verified
"	454600 001F01	7/29/75	791	verified
"	454600 001F01	9/03/75	549	verified
NEW MEXICO	Ms. Cecelia Williams (505) 827 - 5271 Extension 381 Air Quality Division Environmental Improvement Agency			
Albuquerque	320040 003H01	2/28/74	214	verified
"	320040 003H01	7/04/74	216	verified
"	320040 003H01	11/08/75	209	verified
"	320040 003H01	11/14/75	333	verified
Dona Ana	320340 001F02	2/20/74	397	verified
"	320340 001F02	4/10/74	364	verified
Las Cruces	320580 001F01	12/03/74	307	verified
"	320340 001F01	11/14/75	503	changed to 228
"	320340 001F01	11/26/75	340	verified
Raton	320900 001F01	4/11/74	144	verified
"	320900 001F01	10/09/75	185	verified

TABLE 7. CHECK OF TSP DATA FLAGGED AS OUTLIERS BY TELEPHONE INQUIRY (ALL TSP VALUES IN $\mu g/m^3$) (Continued)

Site Name	SAROAD Code	Sampling Date	TSP Value	Remarks
OKLAHOMA	Mr. Scott Thomas (405) 271 - 5220 Air Quality Service Oklahoma State Department of Health			
Muskogee	371980 162F01	7/02/74	235	verified
Tulsa	373000 110F01	4/10/74	468	verified
Tulsa	373000 112F01	3/13/73	479	verified
"	373000 112F01	1/18/75	561	changed to 186
ARKANSAS	Mr. Wilson Tolefree (501) 371 - 1136 Control and Ecology Arkansas Department of Pollution			
Forrest City	040900 001F01	2/21/73	374	verified
"	040900 001F01	6/09/73	401	verified
Helena	041080 001F01	5/17/74	389	verified
Jonesboro	041320 001F01	12/13/74	219	verified
Stuttgart	042540 005F05	8/26/73	512	verified
"	042540 005F05	10/19/73	924	verified

The corrections to the Las Cruces and Tulsa data improve the consistency of the cumulative frequency plots on both lognormal and exponential paper. The new frequency distributions are presented in Figures 15a and 15b for Las Cruces, and in Figures 16a and 16b for Tulsa. The highest reported value at Tulsa (112F01), 479 $\mu g/m^3$ on 3-13-73, still appears to be an outlier on the plots. However, this value is verified by high values at two nearby stations (Muskogee and Tulsa 110F01) on the same date.

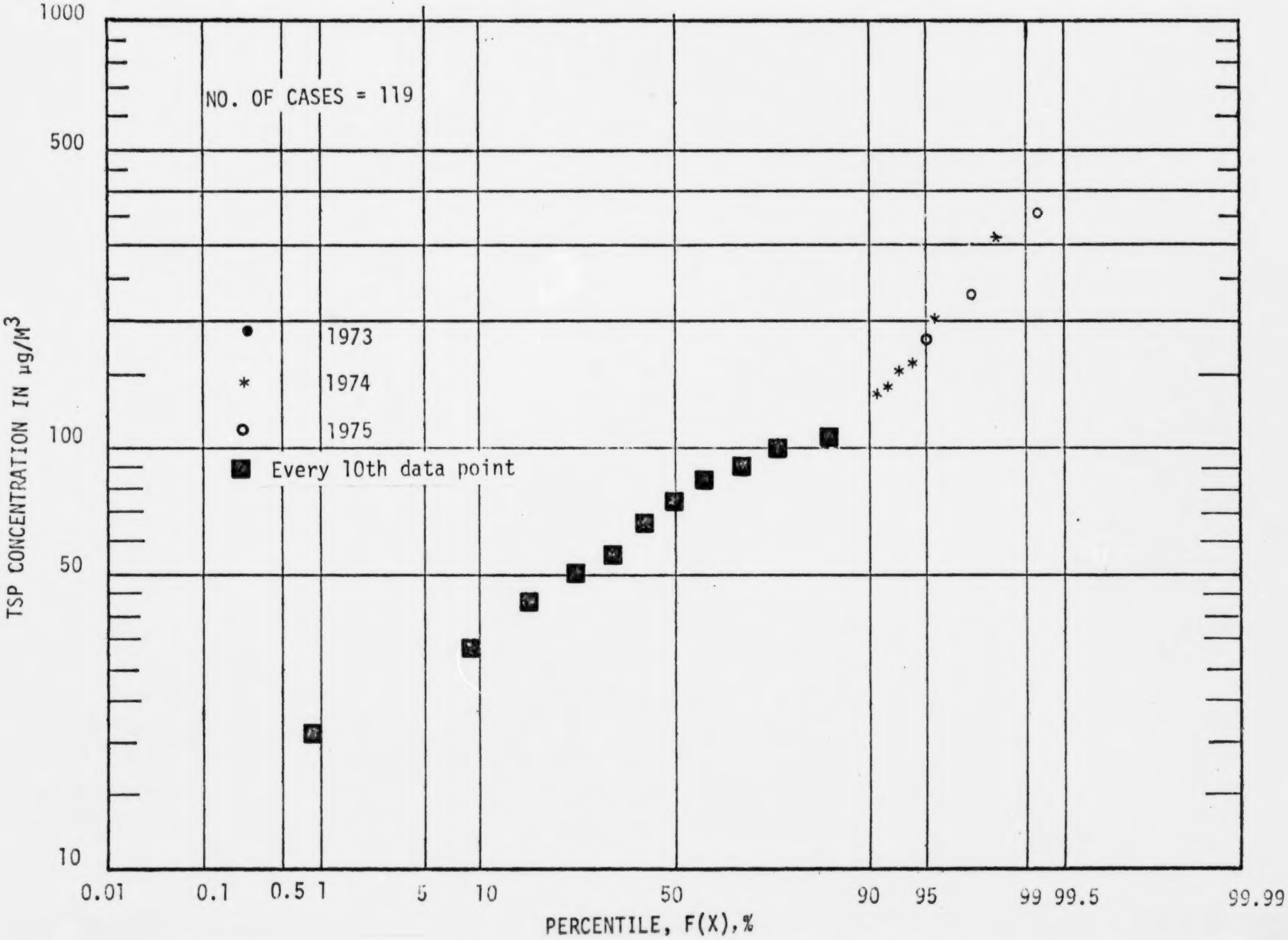


Figure 15a. Corrected log-normal distribution plot of TSP data at Las Cruces, NM (#320580001F01)

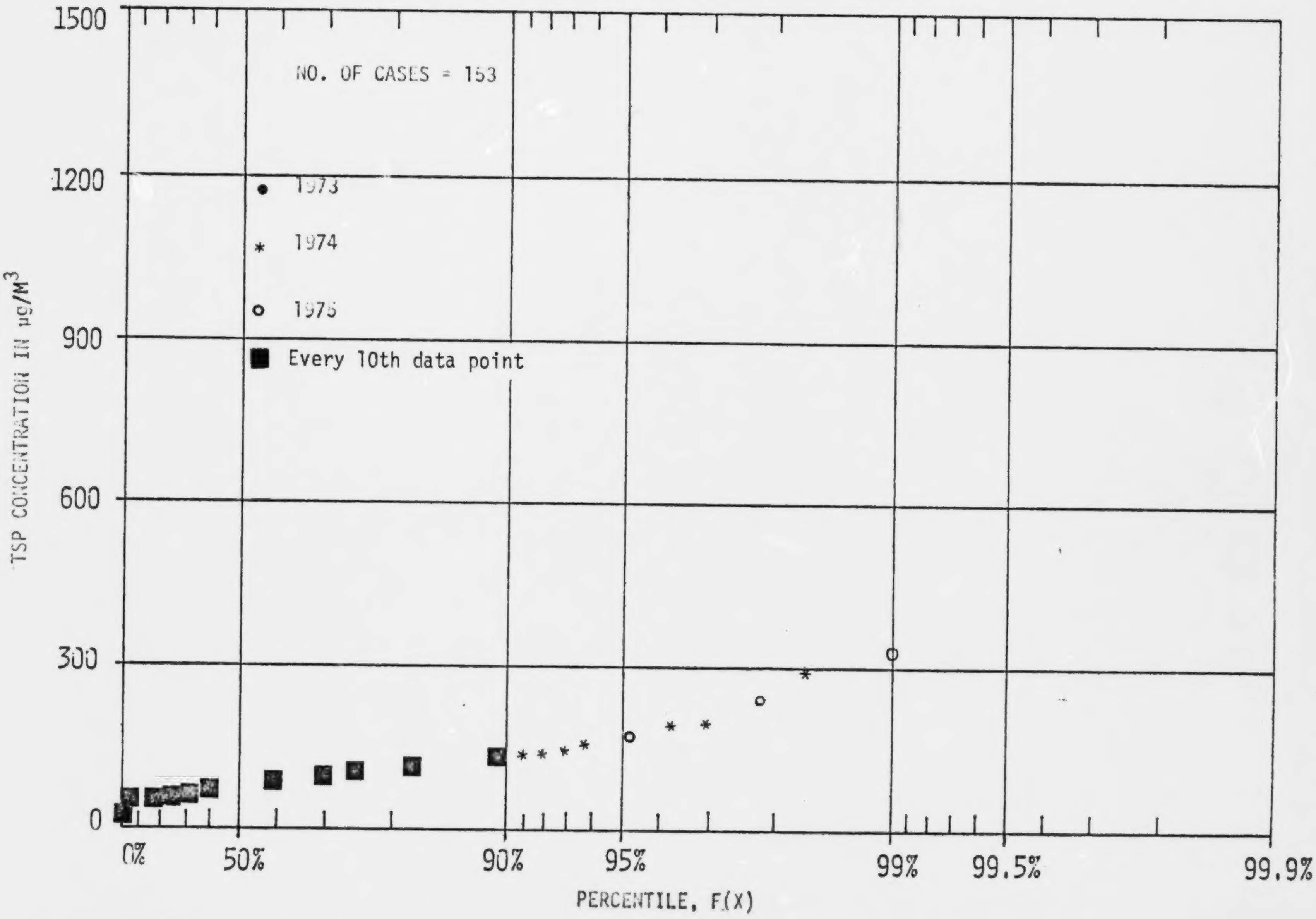


Figure 15b. Corrected exponential distribution plot of TSP data at Las Cruces, NM (#320580001F01)

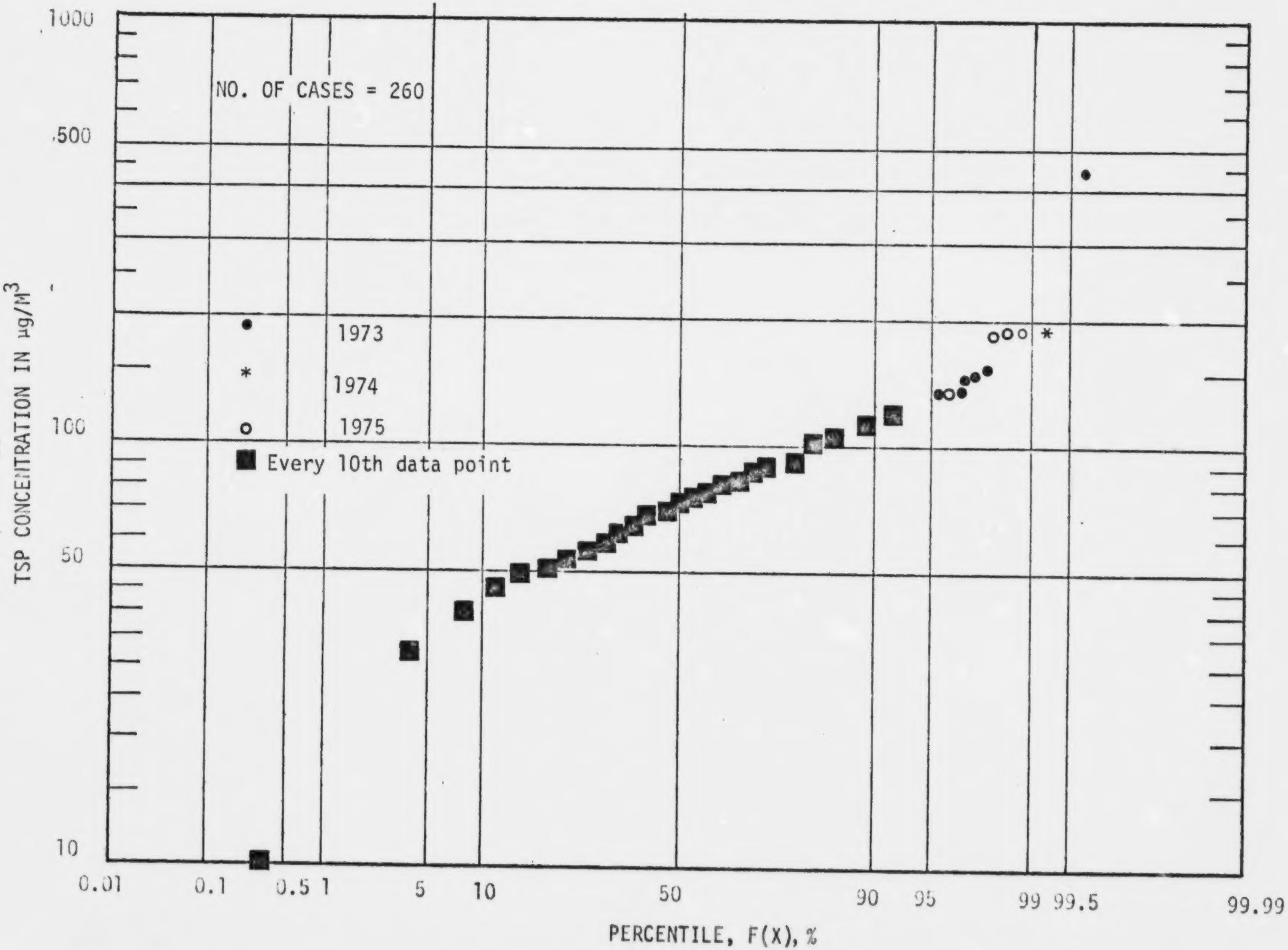


Figure 16a. Corrected log-normal distribution plot of TSP data at Tulsa, OK (#373000112F01)

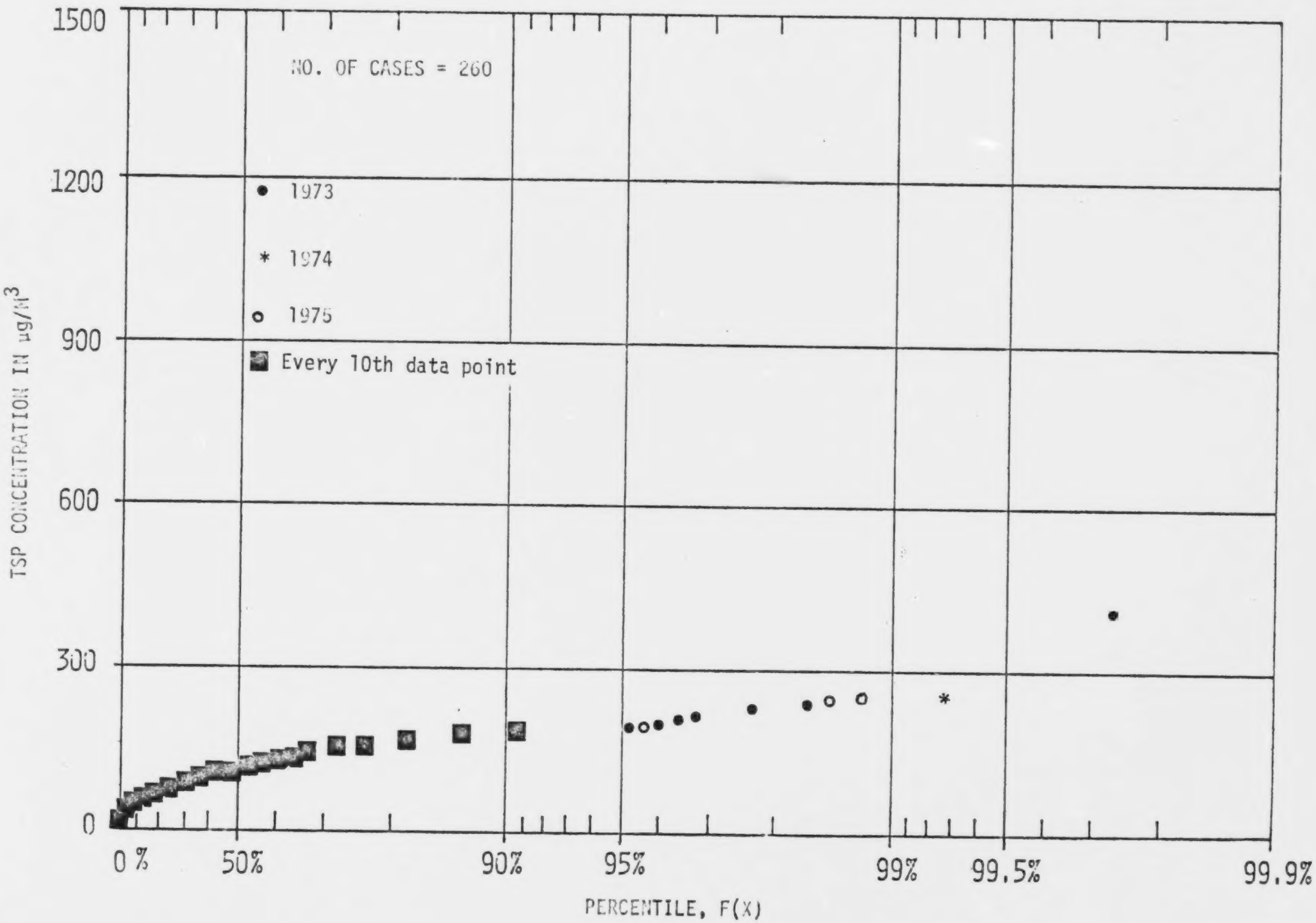


Figure 16b. Corrected exponential distribution plot of TSP data at Tulsa, OK (#373000112F01)

4.0 DETERMINATION OF STATISTICAL PARAMETERS FOR THE TSP DATA

To provide familiarity with the data base and to serve as a reference for later chapters, this chapter summarizes some of the basic statistics for the TSP data. This chapter also includes a brief discussion of the cumulative frequency distributions of TSP data.

4.1 SUMMARY STATISTICS OF TSP DATA

Table 8 presents statistics of the TSP data for the 25 locations under study. The table includes the following:

- Number of measurements in 1973, 74, and 75.
- Arithmetic mean, geometric mean, and geometric standard deviation based on the entire 3-year record.
- The highest and second-highest measurements each year.
- Expected highest and expected second-highest values for annual sample sizes of 60 and 365 days.

The expected highest and expected second-highest values are obtained by drawing a straight line that fits the upper-end of the cumulative frequency distributions plotted on exponential paper. Figure 17 presents an example.* Approximately the same results are obtained if one fits the upper-end of the cumulative frequency distribution with a straight line on log-normal paper.

Table 8 demonstrates that, for most study locations, the geometric mean concentrations range from approximately 40 to 100 $\mu g/m^3$, with the arithmetic means about 10 - 20% higher. The expected maxima are typically around 150 to 350 $\mu g/m^3$ based on 60 samples per year and around 200 to 600 $\mu g/m^3$ based on 365 samples per year. The two major exceptions to the above observations are Stuttgart (005) and San Benito which record TSP concentrations about twice the average level of the other sites. All of

---

*The percentile for the rth highest value in a sample of size N is $P = (1 - \frac{r}{N + 1}) \times 100\%$. Thus, the percentiles for the expected highest and second-highest values in a sample size of 60 are 98.4% and 96.7%, respectively. For a sample size of 365, the expected highest and second-highest values correspond to percentiles of 99.73% and 99.45%, respectively.

TABLE 8. SUMMARY STATISTICS OF TSP DATA IN THE STUDY AREA

SITE NAME I.D. NUMBER	SAMPLE SIZE 73	SAMPLE SIZE 74	SAMPLE SIZE 75	ENTIRE RECORD AM*	ENTIRE RECORD GM*	ENTIRE RECORD GSD	73 HIGHEST* / 2ND HIGHEST	74 HIGHEST* / 2ND HIGHEST	75 HIGHEST* / 2ND HIGHEST	EXPECTED ANNUAL MAX*/2ND MAX. N = 60	EXPECTED ANNUAL MAX*/2ND MAX. N = 365
Forrest City 040900001F01	58	59	55	98.1	87.2	1.63	401/374	188/174	223/174	260/220	370/330
Helena 041080001F01	59	61	54	91.6	79.7	1.73	214/208	389/242	169/160	250/220	350/310
Jonesboro 041320001F01	56	57	59	79.3	71.4	1.60	177/146	219/218	147/133	200/180	260/230
Springdale 042480001F01	58	58	54	76.3	70.2	1.53	143/133	181/158	122/118	170/150	200/190
Stuttgart 042540001F01	60	61	53	103.9	91.5	1.66	223/192	399/234	450/226	290/250	380/350
Stuttgart 042540005F05	30	52	49	189.9	159.7	1.84	924/512	390/330	475/357	600/500	900/800
Albuquerque 320040003H01	0	45	54	114.4	101.1	1.74	M	216/214	333/209	270/230	340/300
Albuquerque 320040008H02	4	42	53	81.2	68.3	2.00	67/58	166/159	194/176	180/170	210/200
Dona Ana Co. 320340001F02	9	86	47	100.0	80.2	2.03	296/114	397/364	255/213	310/270	410/370
Las Cruces 320580001F01	0	75	44	84.5	71.8	1.72	M	307/229	340/175	300/230	430/370
Raton 320900001F01	0	43	46	50.5	44.8	1.63	M	144/107	185/105	150/120	220/200
Bethany 370260014F01	96	82	45	83.0	73.0	1.68	255/225	238/219	170/119	220/190	300/270
Muskogee 371980162F01	94	91	80	57.2	50.6	1.68	150/118	235/159	126/119	160/120	240/210

TABLE 8. SUMMARY STATISTICS OF TSP DATA IN THE STUDY AREA (Con't)

SITE NAME I.D. NUMBER	SAMPLE SIZE			ENTIRE RECORD			73 HIGHEST* 2ND HIGHEST	74 HIGHEST* 2ND HIGHEST	75 HIGHEST* 2ND HIGHEST	EXPECTED ANNUAL MAX.*/2ND MAX.	
	73	74	75	AM*	GM*	GSD				N = 60	N = 365
Oklahoma City 372200002F01	98	67	0	46.4	39.1	1.79	252/137	214/204	M	200/150	290/250
Roger Mills 372600900F01	33	37	43	80.5	67.8	1.74	314/245	416/347	170/165	340/250	540/470
Sequoyah Co. 372760471F01	63	41	37	47.3	38.0	1.97	176/154	134/130	148/122	160/140	210/200
Tulsa 373000110F01	84	112	47	78.1	65.2	1.77	365/335	468/285	260/158	310/250	460/400
Tulsa 373000112F01	109	110	41	78.7	70.6	1.58	479/180	188/125	186/136	200/170	300/280
Corpus Christi 451150001F01	44	40	50	81.6	75.8	1.47	154/118	167/149	265/147	190/170	240/220
Corpus Christi 451150004F02	31	48	49	109.9	95.3	1.72	334/165	291/275	332/246	310/280	420/380
Corpus Christi 451150019F01	16	46	49	73.5	68.5	1.45	103/93	145/127	201/149	170/160	220/200
Harlingen 452320001F01	59	50	47	99.2	85.0	1.78	143/141	298/264	231/165	280/240	360/320
Lubbock 453340001F01	50	32	32	97.5	83.6	1.70	339/165	340/289	496/268	380/300	600/500
McAllen 453390001F01	51	46	40	105.1	95.0	1.59	320/221	291/224	195/190	280/230	370/330
San Benito 454600001F01	49	42	49	146.0	111.9	1.94	1302/424	286/258	883/791	850/550	1450/1250

*Units are in $\mu g/m^3$

M : TSP data are missing in that year

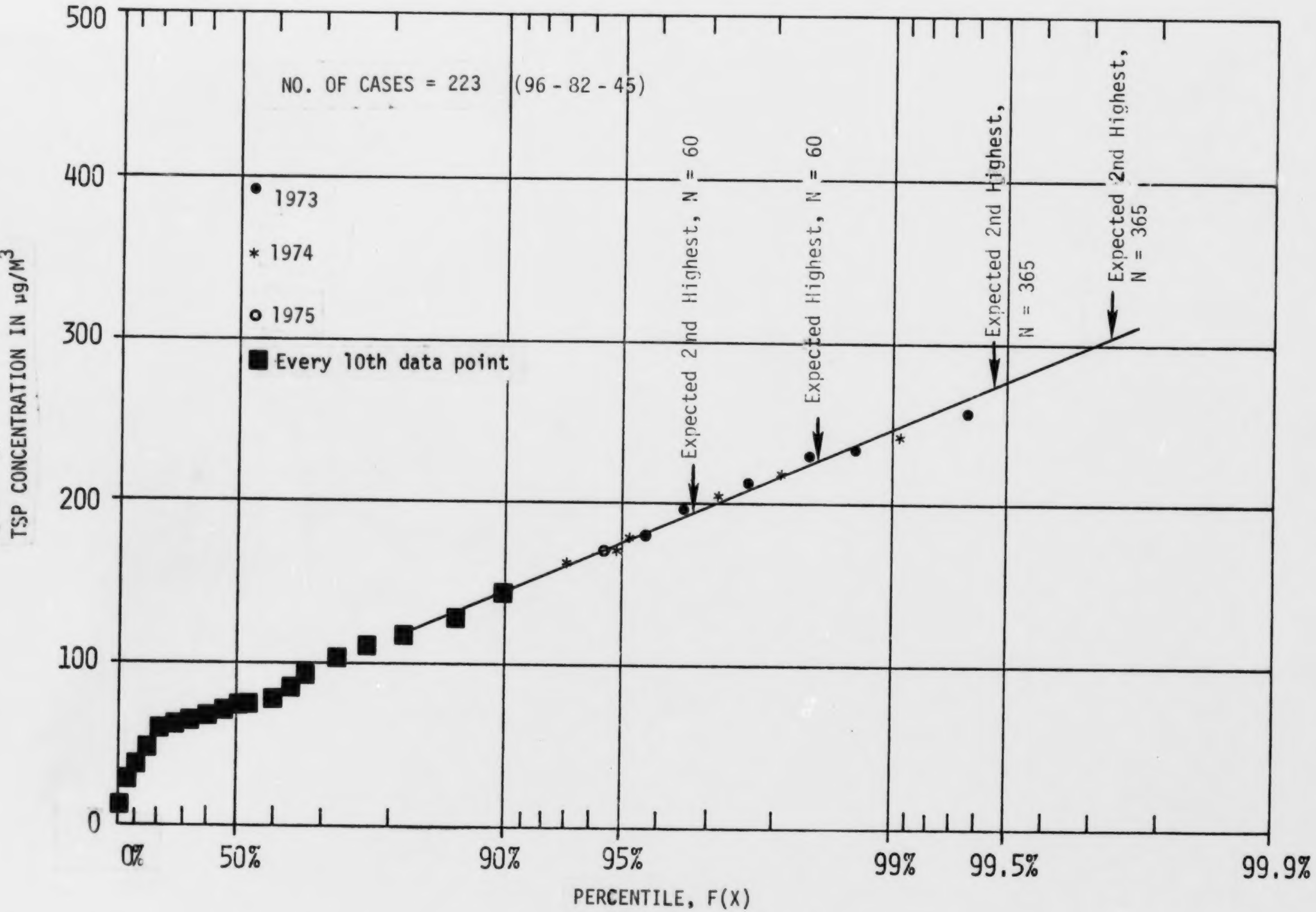


Figure 17. Exponential distribution plot of TSP data at Bethany, OK (#370260014F01)

the sites show fairly high TSP concentrations because we have selected locations that have been considered nonattainment sites by EPA.

We have computed expected highest and second-highest values for both 60 samples per year (as commonly practiced) and 365 samples per year (every day sampling). Table 9, derived from the results listed in Table 8, indicates that the attainment status of the 25 sites for the short-term TSP standards is sensitive to the choice of sampling frequencies.

TABLE 9. SENSITIVITY OF ATTAINMENT STATUS TO YEARLY SAMPLING FREQUENCY

STANDARD	NUMBER OF SITES EXCEEDING STANDARD			
	Expected Annual Maximum		Expected Annual 2nd Maximum	
	N = 60	N = 365	N = 60	N = 365
Primary (260 $\mu g/m^3$)	13	18	5	16
Secondary (150 $\mu g/m^3$)	25	25	22	25

4.2 CUMULATIVE FREQUENCY DISTRIBUTIONS

Appendix A presents cumulative frequency distributions for the TSP data at the 25 sites, plotted on both lognormal and exponential paper.* Overall, most of the sites fit a log-normal distribution (straight line on lognormal paper) quite well. The extreme upper tail of the distribution, plotted on lognormal paper, often deviates somewhat from the remainder of the distribution; this indicates that a separate lognormal distribution should be fit to the tail in order to predict the extreme upper percentiles.

The exponential distribution (straight line on exponential paper) usually does not fit the main body and lower tail of the data very well. However, the exponential distribution usually fits well to the upper tail of the distribution. The plots on exponential paper are particularly useful for examining the extreme values because exponential paper emphasizes the high end of the concentration distribution.

---

*Revised plots for Las Cruces and Tulsa (112F01), incorporating changes resulting from the data quality check, are presented in Figures 15 and 16.

Chapter 8 of this report will show how time series of TSP data can be analyzed to determine if standard violations may have been caused by temporary activity, such as construction. Some clues concerning the possible importance of temporary activity can also be gained by examining the cumulative frequency distributions. In Appendix A individual points, distinguished by year (1973-1975), are plotted for the upper few percentiles. If nearly all the high values occur in a single year, the high values might have been caused by temporary activity.[*] In this regard we make the following notes concerning high values concentrated in certain years:

- Years of very high values
  Harlingen, Texas - 1974

- Years of high values
  Forrest City, Arkansas - 1973
  Springdale, Arkansas - 1974
  Muskogee, Oklahoma - 1974

- Years of moderately high values
  Jonesboro, Arkansas - 1974
  Stuttgart (005), Arkansas - 1973
  Albuquerque (008), New Mexico - 1975
  Dona Ana, New Mexico - 1974
  Corpus Christi (001), Texas - 1975
  Corpus Christi (019), Texas - 1975

---

[*] The occurrence of high values in 1973 might also reflect an effective control policy beginning in 1974. The occurrence of high values in 1975 might reflect new emission sources.

5.0 DECISION-TREE ANALYSIS OF TSP AND METEOROLOGICAL DATA

The primary goal of this project is to identify sources of high TSP concentrations by studying the relationship between daily TSP levels and meteorology. In order to characterize the relationship between TSP and meteorology, we have chosen to use the AID decision-tree program developed at the University of Michigan Institute for Social Research (Sonquist et al 1973). The AID program simulates the procedures of a good researcher in searching for predictors that best account for the variations of a dependent variable. AID sequentially splits the data among prestated ranges in the independent variables, at each stage choosing the split that maximizes the variance explained in the dependent variable. The result is a decision-tree that accounts for the variance in the dependent variable (TSP) according to groups (meteorological classes) defined by ranges in the independent variables.

For the problem at hand, AID offers several advantages over more conventional data-analytic techniques such as multiple linear regression. These advantages include the following:

- The AID program is based on a very general form of the least-squares principle which does not involve restrictive assumptions such as additivity and linearity.

- The meteorological classes defined by AID can be interpreted on physical grounds more easily than regression coefficients relating TSP to meteorology.

- The meteorological classes produced by AID are ideal for meteorological normalization of TSP trends (see Chapter 8).

- The AID program generally explains more of the variance in the TSP data than does multiple linear regression or multiple log-linear regression.

This chapter discusses the AID program and describes its application to TSP and meteorological data from 25 locations in EPA Region 6. Also included are comparisons of the AID results for the 25 sites with results obtained using multiple linear regression and multiple log-linear regression.

5.1 THE AID DECISION-TREE PROGRAM

The purpose of the AID program is to illustrate the relationship between a dependent variable A and independent variables x, y, z, ... by dividing the available data into classes (defined by the independent variables) in such a way as to best explain the variance in A. The data base for AID consists of a set of "n" measurements: $A_1, \ldots, A_n$ and $x_1, \ldots x_n$; $y_1 \ldots, y_n$; $z_1 \ldots, z_n$; etc. Each of the independent variables is represented by discrete numbers, e.g. x might assume values of 1, 2, 3, or 4; y might assume values of 1, 2, 3, 4, 5, 6, or 7; z might assume values of 1, 2, 3, 4, or 5; etc. To achieve discreteness, the raw data for the independent variables are usually divided in ranges. For example, if the variable z represents wind speed (WS), one might define $z = 1$ if $0 \leq WS < 3$ mph; $z = 2$ if $3 \leq WS < 7$ mph; $z = 3$ if $7 \leq WS < 15$ mph; $z = 4$ if $15 \leq WS < 25$ mph; and $z = 5$ if $25 \text{ mph} \leq WS$.

The first step of AID determines the independent variable that can be split into two groups (say z = 1 and 4 versus z = 2, 3, and 5) which maximize the residual sum of squares (RSS) explained in the dependent variable. If the two groups (G_1 and G_2) have group means $\overline{A1}$ and $\overline{A2}$, then the RSS explained by splitting the data is

$$\Delta RSS = RSS_0 - RSS_1 - RSS_2$$

where

$$RSS_0 = \sum_{j=1}^{n} (A_j - \overline{A})^2 = \text{RSS for entire data set}$$

$$RSS_1 = \sum_{G_1} (A_j - \overline{A1})^2 = \text{RSS for } G_1$$

$$RSS_2 = \sum_{G_2} (A_j - \overline{A2})^2 = \text{RSS for } G_2$$

The program considers all possible partitions of the data among all the independent variables and selects the specific independent variable and specific partition that maximizes ΔRSS.

The partitioning process is repeated on each of the groups G_1 and G_2 to yield subgroups. These and all further subgroups are divided until either (1) no subgroup can be split to achieve a ΔRSS above a preset lower bound, (2) further division will produce a subgroup with number of elements less than a preset bound, or (3) the number of terminal subgroups reaches a prespecified limit.

In order to save on computer costs, a simplified version of AID is applied in this study. The simplification is that all independent variables are put into discrete forms representing monotonic intervals (as in the wind speed example above), and only monotonic splits of the data are considered (i.e. $z = 1$ versus $z = 2, 3, 4$, and 5; $z = 1$ and 2 versus $z = 3, 4$, and 5; etc.). It should be noted, however, that this process does not limit us to monotonic relationships between the dependent and independent variables. For instance, in Figure 18, a nonmonotonic relationship would be discerned by AID by first splitting the data according to $z = 1$ and 2 versus $z = 3$, 4, and 5, and then splitting $z = 3$ and 4 versus $z = 5$.

5.2 APPLICATION OF "AID" TO TSP AND METEOROLOGICAL DATA

This section discusses the application of AID to TSP and meteorological data at the 25 study locations. The data for the analysis covers the three year period 1973-1975. The reader is referred to Chapter 2 (e.g. Table 1) for a description of the 25 TSP monitoring sites and their pairings with surface and upper-air weather sites.

5.2.1 Meteorological Variables

The data base for each of the 25 locations includes TSP measurements and corresponding values of the 18 meteorological parameters listed in Table 10. The meteorological parameters are extracted or computed from weather data tapes obtained from the National Climatic Center (NCC) for 10 surface monitoring sites and 6 upper-air monitoring sites. Six of the weather variables (AVVIS, NBLO, AVWS, AVWD, AVRH and WV) are obtained from surface

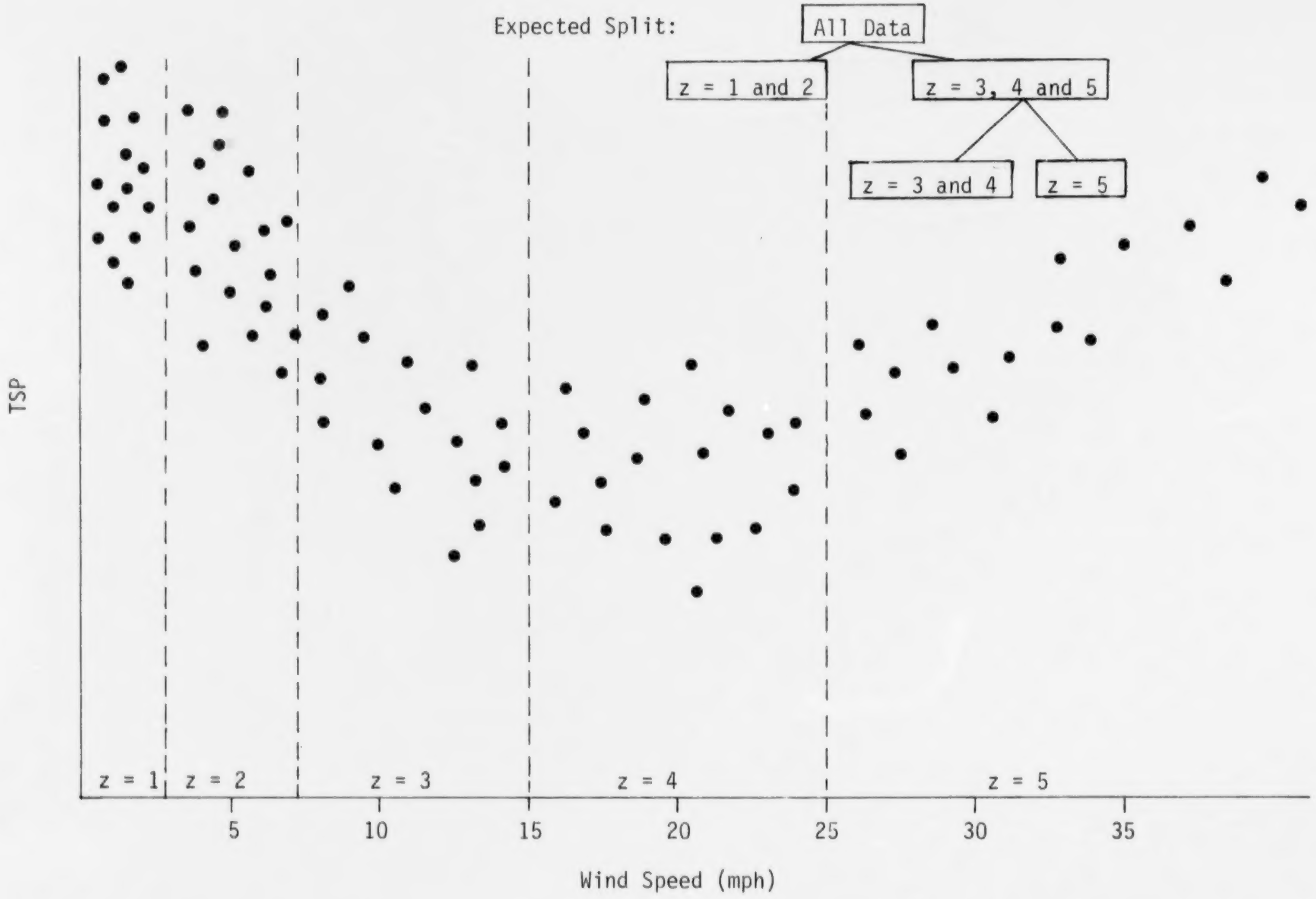


Figure 18. Hypothetical example of nonmonotonic relationship that would be discerned by AID.

TABLE 10. SUMMARY OF METEOROLOGICAL VARIABLES

			RANGES FOR DIGITIZATION OF THE VARIABLE*								
VARIABLE	CODE	UNITS	1	2	3	4	5	6	7	8	9
Month of the Year	MON	---	[1,2]	[3,4]	[5,6]	[7,8]	[9,10]	[11,12]	---	---	---
Daytime Average Visibility	AVVIS	miles	[0,1.5)	[1.5,3)	[3,5)	[5,9)	[9,14)	[14,25)	[25,40)	[40,60)	[60,∞]
Number Observations Blowing Dust	NBLO	---	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	---
Arithmetic Average Wind Speed	AVWS	knots	[0,2)	[2,3)	[3,5)	[5,7)	[7,10)	[10,15)	[15,25)	[25,35)	[35,∞]
Vector Average Wind Direction	AVWD	degrees from N	[0,45)	[45,90)	[90,135)	[135,180)	[180,225)	[225,270)	[270,315)	[315,360)	---
Average Relative Humidity	AVRH	percent	[0,30)	[30,40)	[40,50)	[50,60)	[60,65)	[65,70)	[70,80)	[80,90)	[90,100]
Wind Variability	WV	percent	[0,2)	[2,5)	[5,10)	[10,20)	[20,30)	[30,40)	[40,60)	[60,80)	[80,100]
A.M. Mixing Height	AMMH	meters	[0,50)	[50,100)	[100,200)	[200,300)	[300,500)	[500,750)	[750,1000)	[1000,1500)	[1500,∞]
A.M. Average Wind Speed Through Mixing Layer	AMWLY	meters/sec.	[0,1)	[1,2)	[2,3.5)	[3.5,5)	[5,7.5)	[7.5,10)	[10,15)	[15,20)	[20,∞]
P.M. Mixing Height	PMMH	meters	[0,200)	[200,300)	[300,500)	[500,750)	[750,1000)	[1000,1500)	[1500,2000)	[2000,2500)	[2500,∞]
P.M. Average Wind Speed Through Mixing Layer	PMWLY	meters/sec.	[0,1)	[1,2)	[2,3.5)	[3.5,5)	[5,7.5)	[7.5,10)	[10,15)	[15,20)	[20,∞]
Maximum Temperature	MXTMP	°F	[-∞,40)	[40,50)	[50,60)	[60,70)	[70,80)	[80,90)	[90,100)	[100,110)	[110,∞]
Minimum Temperature	MNTMP	°F	[-∞,10)	[10,20)	[20,30)	[30,40)	[40,50)	[50,60)	[60,70)	[70,80)	[80,∞]
Amount of 1-Day Precipitation	PREC1	inches	[0,.01)	[.01,.1)	[.1,.25)	[.25,.5)	[.5,1.0)	[1.0,1.5)	[1.5,2.0)	[2.0,3.5)	[3.5,∞]
Maximum Wind Speed	WSMAX	mph	[0,5)	[5,10)	[10,15)	[15,20)	[20,30)	[30,40)	[40,50)	[50,60)	[60,∞]
3-Day Accumulated Precipitation	PREC3	inches	[0,.01)	[.01,.1)	[.1,.25)	[.25,.5)	[.5,1.0)	[1.0,2.0)	[2.0,3.5)	[3.5,5.0)	[5.0,∞]
Number of Days Since Last Precipitation or Snow Cover	ND	days	[0]	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Number of 1-Day Precipitation Observations	NPREC	---	[0]	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]

* "[" implies endpoint is included in the range; ")" implies endpoint is excluded.

weather tapes (NCC - TDF 14); four of the weather variables (AMMH, AMWLY, PMMH, and PMWLY) are obtained from mixing height tapes; and seven of the weather variables (MXTMP, MNTMP, PREC1, WSMAX, PREC3, ND, and NPREC) are obtained from climatic summary tapes (NCC - DECK 345). The variable "month" is, of course, listed on all the TSP and weather tapes.

With the exception of PREC3 (3 day accumulated precipitation) and ND (number of days since last precipitation or snow cover), all of the variables pertain only to data taken on the same day as the TSP measurement. The definitions of all the variables are fairly straight forward, with the exception of WV (wind variability). WV is defined by

$$WV = 100\% \left(1 - \frac{VAVWS}{AVWS}\right)$$

where VAVWS is the daily vector-average wind speed and AVWS is the daily arithmetic-average wind speed.

Table 10 lists the discrete intervals that are used to digitize the meteorological parameters; as noted earlier, AID requires the independent variables to be input in digitized form. In order to allow greater flexibility in the analysis, the actual values of the meteorological variables are maintained in the data base, with digitization performed as the data are input to AID.

5.2.2 Results of AID

The AID program is applied using three termination criteria: (1) the ΔRSS for each step must exceed 1% of the original RSS_0 for the TSP concentrations, (2) no group of less than 3 data points is allowed, and (3) a maximum of ten terminal nodes is allowed. The decision-trees produced by AID are presented in Appendix B for all 25 locations.* Figure 19 provides an example, the decision-tree for Forrest City, AR.

---

* The reader might note that the total sample sizes for the decision-tree analysis are slightly smaller than the number of available TSP samples (as listed in Chapters 2 through 4). This is because days with one or more missing meteorological parameter have been eliminated from the decision-tree analysis.

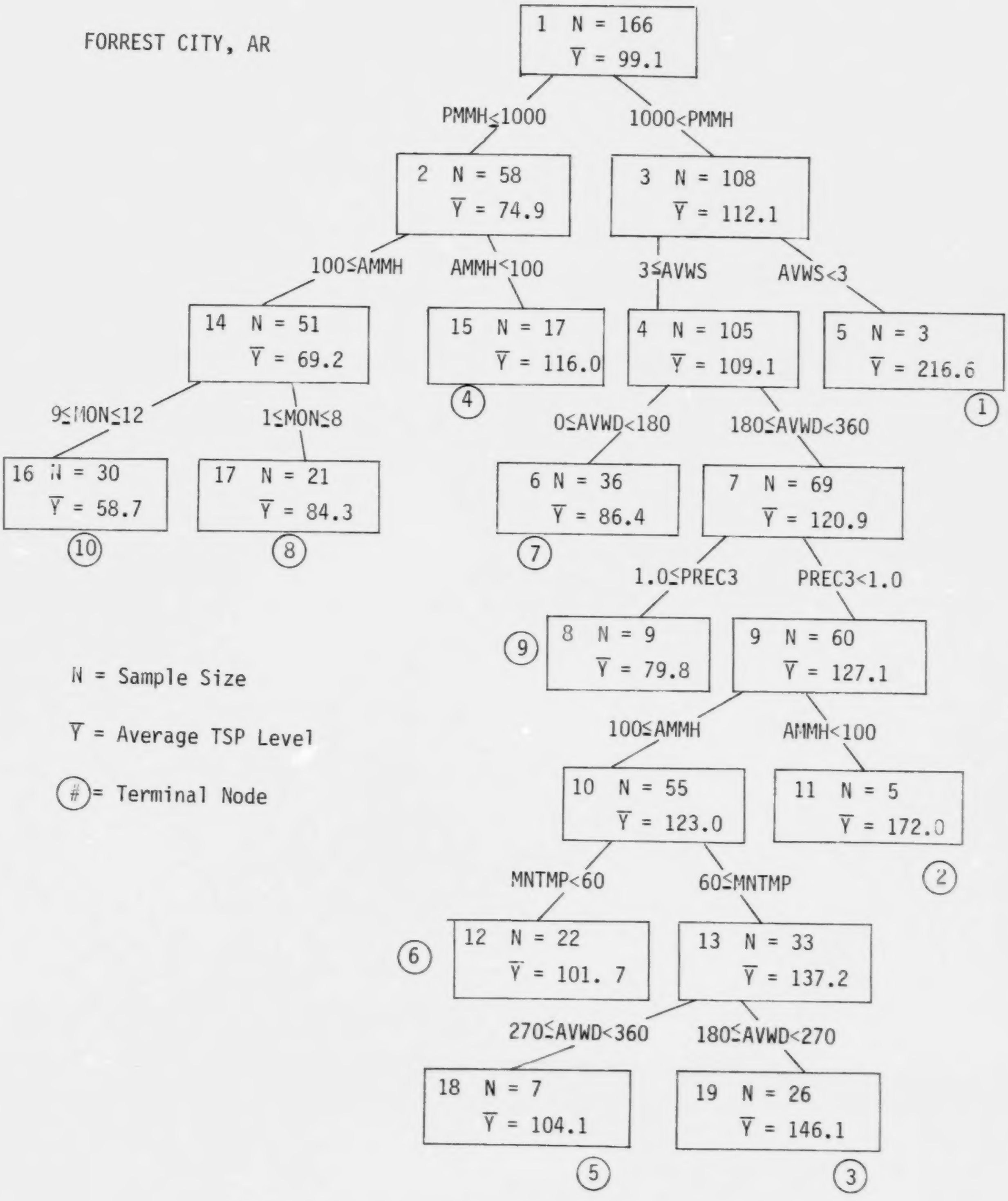

FORREST CITY, AR
1 N = 166
$\overline{Y}$ = 99.1
PMMH≤1000
1000<PMMH
2 N = 58
$\overline{Y}$ = 74.9
3 N = 108
$\overline{Y}$ = 112.1
100≤AMMH
AMMH<100
3≤AVWS
AVWS<3
14 N = 51
$\overline{Y}$ = 69.2
15 N = 17
$\overline{Y}$ = 116.0
④
4 N = 105
$\overline{Y}$ = 109.1
5 N = 3
$\overline{Y}$ = 216.6
①
9≤MON≤12
1≤MON≤8
0≤AVWD<180
180≤AVWD<360
16 N = 30
$\overline{Y}$ = 58.7
⑩
17 N = 21
$\overline{Y}$ = 84.3
⑧
6 N = 36
$\overline{Y}$ = 86.4
⑦
7 N = 69
$\overline{Y}$ = 120.9
1.0≤PREC3
PREC3<1.0
⑨
8 N = 9
$\overline{Y}$ = 79.8
9 N = 60
$\overline{Y}$ = 127.1
N = Sample Size
$\overline{Y}$ = Average TSP Level
(#) = Terminal Node
100≤AMMH
AMMH<100
10 N = 55
$\overline{Y}$ = 123.0
11 N = 5
$\overline{Y}$ = 172.0
②
MNTMP<60
60≤MNTMP
⑥
12 N = 22
$\overline{Y}$ = 101. 7
13 N = 33
$\overline{Y}$ = 137.2
270≤AVWD<360
180≤AVWD<270
18 N = 7
$\overline{Y}$ = 104.1
⑤
19 N = 26
$\overline{Y}$ = 146.1
③

Figure 19. AID decision-tree for Forrest City, AR.

The decision-trees for 23 of the sites contain 10 terminal nodes, the maximum allowable. At one site (Roger Mills Co., OK), the tree contains 9 final nodes; at another site (Lubbock, TX), the tree contains 8 final nodes. The percent variance explained by AID ranges from 28% to 72% and averages 51% over the 25 locations; these are equivalent to total correlation coefficients ranging from 0.53 to 0.85 and averaging 0.71.

Table 11 summarizes the variables selected by AID at each of the 25 locations (multiple asterisks in Table 11 indicate that the variable is used more than once in a decision-tree). As indicated by the summary statistics at the bottom of Table 11, the most frequently used individual variables (approximately in order of importance) are MON, ND, PPMH, PREC3, WV, AVRH, AMMH, MXTMP, AVWD, PREC1, AVWS, WSMAX, PMWLY, MNTMP, AMWLY, and AVVIS.

The four wind speed parameters (AVWS, AMWLY, PMWLY, and WSMAX) are very important; at least one of them is selected at every site but Oklahoma City. Wind variability (WV) and wind direction (AVWD) are also significant, being used at 14 and 12 sites, respectively. Three of the precipitation variables (PREC1, PREC3, and ND) are quite prevalent; at least one of them appears in the decision-tree for 20 of the sites. The mixing height variables are also important; AMMH and/or PMMH are selected at 20 locations. Either or both of the temperature parameters (MXTMP and MNTMP) appear at 19 locations. Other significant variables include month (appearing at 16 sites), relative humidity (13 sites), and visibility (7 sites).

Considering that dust is known to be an important contribution to TSP in Region 6 , it seems very remarkable that the number of daily blowing dust observations (NBLO) occurs in only one decision-tree (at Lubbock, TX). Upon examing the data in detail, the reason for this paradox becomes very obvious. The variable NBLO was always zero at all locations except three: Lubbock, with 4 days of blowing dust observations; Las Cruces, with 2 days; and Dona Ana, with 1 day. Since the AID criteria required a minimum group size of 3, Lubbock was the only location where it was possible for NBLO to appear in the decision-tree.

It should be noted that the sparsity of blowing dust observations does not necessarily imply that actual occurrences of blowing dust are rare. Airport observers are required to report blowing dust only when it re-

TABLE 11. SUMMARY OF METEOROLOGICAL VARIABLES SELECTED BY AID

VARIABLES SELECTED BY AID (Number of asterisks refers to number of splits according to that variable)

SITE	MON	AVVIS	NBLO	AVWS	AVWD	AVRH	WV	AMMH	AMWLY	PMMH	PMWLY	MXTMP	MNTMP	PREC1	WSMAX	PREC3	ND	NPREC
Forrest City, AR	*			*	**			**		*			*			*		
Helena, AR					**	*		*	*		*		*			*	*	
Jonesboro, AR	*			*				*	*	*		*			**	*		
Springdale, AR	*	*			*	*	*			*				*	*		*	
Stuttgart (#1), AR					**		***			*	*				*	*		
Stuttgart (#2), AR				*			*			**		**			*	**		
Albuquerque (#1), NM	*	*				*	*					*			**		**	
Albuquerque (#2), NM	*				**	*	*	*		*	*		*					
Dona Ana, NM		*			*	**	*	*	*					*			*	
Las Cruces, NM	**			**				**							**		*	
Raton, NM	**			*	**	*	*				*	*						
Bethany, OK	*			*		**					*	*		*			**	
Muskogee, OK	*					*	*	*	*			*			*	*	*	
Oklahoma City, OK	*	*			**	**						*	*			*		
Roger Mills Co., OK						**	*		*	*			*	*			*	
Sequoyah Co., OK						*				*	**	**	*		*	*		
Tulsa (#1), OK	*			***			*	*				*				*	*	
Tulsa (#2), OK	**			*	*		*	*						*		**		
Corpus Christi (#1), TX		*		*		*		*		**	*	*					*	
Corpus Christi (#2), TX					*		*	*	*	**	*		**					
Corpus Christi (#3), TX	*	*						*	*	**	*						**	
Harlingen, TX	*			**		*				*		*	*			*	*	
Lubbock, TX		*	*		*		*					*			*		*	
McAllen, TX	*				*					*	*		*		*	*	**	
San Benito, TX	***						*	*	*	*		*	*					
Total Number of Sites at which variable is used	16	7	1	10	12	13	14	13	8	14	10	13	10	5	10	12	14	0

stricts visibility to less than 7 miles, which corresponds to a dust loading of approximately 3000 $\mu g/m^3$ (Orgill and Sehmel 1976). Thus, there may be considerable concentrations of wind-blown dust present without notations of blowing dust being made by the airport meteorologists.

5.3 COMPARISON OF AID RESULTS TO REGRESSION RESULTS

To allow an assessment of AID's performance, regression analyses were conducted for comparison. Multiple linear and multiple log-linear regressions were conducted between TSP and all 18 meteorological variables.

Table 12 compares the percent variance explained by AID to the percent variance explained by multiple regressions. The second column of Table 12 lists the results obtained with the AID criteria selected for use in this report (ΔRSS at least 1%; group size at least 3; no more than 10 terminal nodes). The first column presents results for a relaxed set of AID criteria (ΔRSS at least .5%; group size at least 3; no more than 15 terminal nodes). The last two columns list results for multiple regressions.

Table 12 indicates that AID performs considerably better than multiple linear (or log-linear) regression in explaining TSP levels in terms of meteorological variables. The average variance explained by AID is 51%, compared to 36% for multiple linear regression and 39%* for multiple log-linear regression. The performance of AID relative to regression is especially remarkable because the final decision-trees use only a few of the meteorological parameters, while the regression equations contain all 18 meteorological variables. We could have obtained an even greater average percent variance explained, 59%, by using less stringent criteria in AID, but this would have made the decision-trees more complex and more difficult to interpret.

As evidenced by Table 12, the percent variance explained by AID varies significantly from site to site. One might suspect that the low percent variance explained at some locations is due to large distances between the TSP sampler and the weather monitoring location. This, however, does not appear to be the major reason for low percent variance explained at certain

---

* This actually represents the percent variance explained in ℓn TSP. The percent variance explained in the original TSP variable by the logarithmic regression formula would probably be slightly less.

TABLE 12. PERCENT VARIANCE EXPLAINED BY "AID" COMPARED TO PERCENT VARIANCE EXPLAINED BY MULTIPLE REGRESSIONS AGAINST ALL 18 METEOROLOGICAL VARIABLES

SITE	PERCENT VARIANCE EXPLAINED			
	AID Program		Regression	
	15 Groups	10 Groups	Linear	Log-Linear
Forrest City, AR	50%	42%	31%	44%
Helena, AR	60%	52%	31%	42%
Jonesboro, AR	60%	51%	42%	47%
Springdale, AR	63%	52%	38%	43%
Stuttgart (1), AR	61%	53%	29%	41%
Stuttgart (2), AR	67%	62%	49%	52%
Albuquerque (1), NM	70%	62%	56%	45%
Albuquerque (2), NM	73%	59%	42%	31%
Dona Ana, NM	63%	54%	58%	63%
Las Cruces, NM	62%	56%	33%	50%
Raton, NM	52%	44%	19%	23%
Bethany, OK	51%	39%	31%	41%
Muskogee, OK	36%	28%	24%	29%
Oklahoma City, OK	55%	50%	37%	43%
Roger Mills County, OK	47%	45%	32%	40%
Sequoyah, OK	69%	61%	50%	54%
Tulsa (1), OK	69%	59%	28%	29%
Tulsa (2), OK	44%	40%	24%	13%
Corpus Christi (1), TX	50%	43%	28%	18%
Corpus Christi (2), TX	67%	49%	13%	27%
Corpus Christi (3), TX	63%	54%	40%	28%
Harlingen, TX	71%	61%	60%	57%
Lubbock, TX	76%	72%	66%	53%
McAllen, TX	47%	38%	24%	24%
San Benito, TX	59%	55%	20%	40%
Mean	59%	51%	36%	39%

locations. The 13 sites which are at greatest distances from the surface weather stations[*] have an average percent variance explained of 49.5%, which is nearly as great as the 51% average over all sites.

Table 13 compares the variables selected by AID to the variables that are statistically significant[**] in multiple regressions. The variables used by AID are often the same as the variables that are significant in multiple regressions. For each location, however, the number of variables used by AID is greater than the number of significant variables according to linear or log-linear regression. One possible explanation is that AID discerns significant nonlinear dependencies, whereas the regressions are restricted linear (or log-linear) relationships. Another explanation might be that the AID criteria we have chosen may be equivalent to a less than 95% significance level.

---

[*] These 13 sites are Forrest City, Helena, Jonesboro, Springdale, Stuttgart (#1), Stuttgart (#2), Dona Ana, Las Cruces, Raton, Bethany, Roger Mills, Harlingen, and McAllen. The reader is referred to Figure 2 in Chapter 2 for distances between TSP samplers and weather sites.

[**] For statistical significance in the multiple regressions, we have required that the t-statistic for the regression coefficient exceed 1.9. This criteria corresponds to slightly less than a 95% confidence level.

TABLE 13. SIGNIFICANT METEOROLOGICAL VARIABLES IDENTIFIED BY AID AND BY MULTIPLE REGRESSIONS

	SIGNIFICANT VARIABLES ("A" in AID, "R" in Multiple Linear Regression, "L" in Multiple Log-Linear Regression)																	
SITE	MON	AVVIS	NBLO	AVWS	AVWD	AVRH	WV	AMMH	AMWLY	PMMH	PMWLY	MXTMP	MNTMP	PREC1	WSMAX	PREC3	ND	NPREC
Forrest City, AR	A,R,L			A,R	A	L		A,R	R,L	A		L	A			A,L		L
Helena, AR	R,L				A,R,L	A		A	A,R		A	R,L	A,R			A	A,L	
Jonesboro, AR	A			A				A	A	A,R,L	R,L	A,R,L	R		A	A,L		
Springdale, AR	A,R,L	A			A	A	A			A	R,L	R,L	R,L	A	A,L		A,L	
Stuttgart (#1), AR			R		A,R,L		A			A	A,L			L	A,R,L	A	L	
Stuttgart (#2), AR				A	R,L		A			A,R,L	R	A		R	A	A,R,L		
Albuquerque (#1), NM	A	A	R			A	A,R			R,L		A,R,L	R,L	L	A	L	A,R	
Albuquerque (#2), NM	A,L				A,L	A	A	A		A,R	A,R		A			L		
Dona Ana, NM		A	R,L		A	A,R	A,L	A	A	R		R,L	R,L	A,L	R		A,L	
Las Cruces, NM	A,R,L			A		R,L		A			L				A,L		A	
Raton, NM	A,L			A	A	A,R	A	R			A	A						
Bethany, OK	A		L	A	R	A,R,L		R			A	A,R,L		A			A	
Muskogee, OK	A,L	R,L		L		A,L	A	A	A			A,R,L			A	A	A	
Oklahoma City, OK	A,L	A,R	L		A,L	A,R,L				R		A	A			A,L		
Roger Mills Co., OK				L		A,R	A		A	A		L	A	A,L			A	
Sequoyah Co., OK					R	A,L				A	A,R	A	A,R,L	L	A	A,L		R,L
Tulsa (#1), OK	A			A			A	A			R	A				A,L	A,R	
Tulsa (#2), OK	A,R	R		A	A		A	A						A		A		R,L
Corpus Christi (#1), TX	R	A,R,L		A		A		A		A,R	A	A,R					A,R	
Corpus Christi (#2), TX					A		A	A	A	A	A	L	A				L	
Corpus Christi (#3), TX	A	A,R,L				R		A	A	A	A						A,R,L	
Harlingen, TX	A	L	L	A,R,L		A,R,L		L	R	A,R	R	A	A,L		R,L	A,L	A,R	
Lubbock, TX		A,R	A,R,L	R	A		A					A			A,R,L		A	
McAllen, TX	A,R,L				A	R				A,R	A		A		A	A	A	
San Benito, TX	A		L	L	R,L	R	A,L	A	A	A		A	A,R,L				L	

6.0 INTERPRETATION OF THE DECISION-TREES AND VALIDATION OF THE RESULTS

The previous chapter used the AID decision-tree program to characterize the relationships between TSP and meteorology at the 25 study locations. This chapter examines the TSP-meteorology relationships to gain insight concerning the sources of high ambient TSP levels. The specific goal is to classify causes of high TSP levels according to three types of sources: 1. wind-blown dust (dust raised by the action of winds), 2. man-made dust sources (dust raised by anthropogenic activities), and 3. conventional man-made sources (particulate sources contained in conventional emission inventories, such as particulates in exhaust plumes). The division between the first source category and the latter two source categories is particularly important, because it is often assumed that wind-blown dust is not controllable while the man-made sources are controllable.*

This chapter begins with a general discussion of the relationship between TSP and meteorology and how it can be used to distinguish source types. Next, the AID decision-trees are interpreted and the source types are identified for each of the 25 study locations. The conclusions for several locations are checked against the results of a comprehensive study by the Texas Air Control Board (Price et al 1977). Finally, an evaluation is made of the method as a means to categorize locations as to cause of nonattainment.

6.1 USE OF THE TSP-METEOROLOGY RELATIONSHIP FOR IDENTIFYING SOURCE TYPES

This section presents a general discussion of the relationship between TSP and meteorology and how this relationship can be used to distinguish sources of high TSP levels. The discussion is organized according to various types of weather parameters: precipitation, wind speed, wind direction, wind variability, mixing height, relative humidity, temperature, visibility, month, and observations of blowing dust. We are indebted to Frank Hall, EPA Region 6 Meteorologist, for many of the ideas contained in this section.

---

* Actually, the distinction between wind-blown dust and man-made dust involves a misnomer because wind-blown dust is affected by human activities (agricultural practices, construction, or other activities which govern the condition of the ground cover). Because human activities partly affect wind-blown dust, it is not always reasonable to consider wind-blown dust as "uncontrollable".

Precipitation

The three key precipitation variables in this study are precipitation on the day of the TSP measurement (PREC1), three day total precipitation up to and including the day of the measurement (PREC3), and number of days since last measurable precipitation or last snow cover (ND). The precipitation variables are most useful in distinguishing dust sources (wind-blown or man-made) from conventional man-made sources. If high TSP levels are associated with low values of precipitation (ie. dry ground) and low values of TSP are associated with high values of precipitation (ie. wet ground) then dust sources would be implicated. If high values of TSP are associated with very high values of ND (say $ND>20$) then wind-blown dust, in particular, might be implicated. If there is little precipitation dependence, or if high TSP values occur with wet ground, then conventional sources would likely be important.

A major exception to these rules would be wind-blown dust produced in front of a squall line or before a heavy thunder shower. If these cases were important, wind-blown dust could be positively related to PREC1. However, we would expect there to be dry ground before such an event, ie. PREC3 should not be greater than PREC1.

Wind Speed

The wind speed variables considered in this study are daily average wind speed (AVWS), daily maximum wind speed (WSMAX), and morning and afternoon wind speed averaged up to the mixing height base (AMWLY and PMWLY). The wind speed variables are critical for distinguishing wind-blown dust from man-made sources (dust related or conventional).

If high TSP levels are associated with very high surface winds, say $AVWS\geq15$ knots or $WSMAX\geq30$ knots, then wind-blown dust would definitely be implicated. High TSP associated with high AMWLY or PMWLY could also implicate wind-blown dust because the wind speeds through the mixing layer are correlated with surface wind speeds.

In some cases. transported wind-blown dust would result in a positive relationship between TSP and AMWLY or PMWLY. In these cases, AMWLY or PMWLY would not have to be extremely high because it generally takes lesser wind speeds to support dust aloft than it does to raise dust from the ground.

If high TSP is associated with low wind speeds, then wind-blown dust is unlikely and man-made sources (dust related or conventional) are implicated. This would be confirmed if low TSP were associated with high winds. A rare exception would be wind-blown dust associated with moderate to low wind speeds in cases where the dust was originally raised by high wind speeds at another location.

Wind Direction

If there is a strong dependence of TSP on wind direction, then the high TSP levels are most likely from man-made sources (possibly from a large point source). However, a wind directional dependence could also be associated with wind-blown dust if the wind-blown dust sources (ie. agricultural activities)lie in a certain direction or if high wind speeds tend to be closely correlated with a certain wind direction.

Wind Variability

Wind variability (WV) tends to be negatively correlated with wind speed. Thus, if high TSP levels are associated with high wind variability, anthropogenic sources would be implicated. Conversely, if high TSP levels are associated with low wind variability, wind-blown dust (possibly transported from a certain direction) would be likely.

An exception to the above rules would be point sources or other man-made sources that are located in a distinct direction from the TSP monitor. In these cases, low wind variability could be associated with high TSP levels caused by man-made sources. We expect, in these cases, that a wind directional dependence would also be discerned by AID.

Mixing Height

The two mixing height variables considered in this study are morning mixing height (AMMH) and afternoon mixing height (PMMH). If high TSP levels are associated with high PMMH, dust related sources (man-made or wind-blown) are implicated. This is because high PMMH is associated with unstable atmospheres (which tend to keep dust suspended) and because high PMMH often occurs with dry air (and probably dry ground). If high TSP is associated with moderate to low AMMH and very high PMMH, we would suspect wind-blown dust in particular because such conditions likely represent very unstable atmospheres and high wind speeds. The above conclusions would be reinforced if low TSP levels were associated with low mixing heights.

If high TSP levels occur when mixing heights are low, man-made sources (probably conventional, possibly dust related) would be implicated. Man-made sources, particularly conventional ones, would be likely if very low TSP levels occur when PMMH is high.

Relative Humidity

A negative relationship between average relative humidity (AVRH) and TSP levels would reflect dust sources (man-made or wind-blown). This is because dry air would tend to be associated with dry ground and with high afternoon mixing height (see previous discussions).

If high TSP levels occur when AVRH is high, man-made sources, particularly conventional ones, would be implicated. In this situation, secondary aerosols (such as suflates and nitrates) might also constitute a significant fraction of TSP, because high humidity levels can help to promote their formation.

Temperature

The two temperature parameters included in this study are daily minimum temperature (MNTMP) and daily maximum temperature (MXTMP). If TSP levels are high when MXTMP is high, dust related sources (man-made or wind-blown) would be likely because such conditions likely reflect dry air, high PMMH, and an unstable atmosphere. This conclusion would be reinforced if TSP is low when MXTMP is low.

If TSP is high when temperature (especially MNTMP) is low, man-made sources would be implicated. Low temperatures could reflect wintry conditions with strong nocturnal inversions trapping man-made emissions.

Visibility

Low average visibility (AVVIS) is usually associated with high concentrations of small (.1 to 1 micron) particles and high relative humidity.

Thus, low visibility would typically reflect anthropogenic sources with significant contributions from secondary aerosols (e.g. sulfates or nitrates) which tend to form in the .1 to 1 micron size range, called the accumulation size range. Wind-blown dust can also reduce visibility, but this tends to require very high dust levels such as those produced by a major dust storm.

Month

Relationships between TSP and time of year may just reflect the predominance of certain meteorological conditions during certain seasons. For example, strong winds in Region VI often occur in spring and late winter; high TSP in those seasons may implicate wind-blown dust.

However, because meteorological parameters are explicitly included in the analysis, we would expect AID to select meteorological variables rather than time of year to describe weather related effects. If this is the case, the dependence of TSP on season should represent seasonal changes in source strength. One source type with strong seasonality is agriculture; thus, seasonal splits in the decision-tree may implicate agricultural sources.

Blowing Dust Observations

Because airport observers are only required to report blowing dust when it restricts visibility to less than 7 miles, blowing dust observations are extremely rare. When these observations are recorded, however, they do pinpoint the source of TSP as being wind-blown dust.

6.2 INTERPRETATION OF THE AID DECISION-TREES

In light of the concepts discussed in the previous section, this section reviews the AID decision trees and identifies sources of high TSP levels at the 25 study locations. As part of the review process, some of the terminal nodes in the decision-trees are combined so that the data are split into only 4 or 5 classes. Combining the terminal nodes this way aids in interpreting the trees and leads to larger sample sizes which are more useful for trend analyses (see later discussion, Chapter 8). The terminal nodes are regrouped into classes using two criteria: (1) the nodes included in each class should have similar meteorological specifications so as to facilitate interpretation, and (2) the means of the nodes included in each class should be fairly close so as to preserve most of the percent variance explained by the original tree.

Table 14 illustrates the re-grouping of the decision-tree for Forrest City, AR* and presents the conclusions concerning sources of high TSP concentrations at that location. Similar tables for all 25 study locations are are presented in Appendix C. The conclusions with respect to the causes of high TSP levels at the 25 study sites are summarized in Table 15. As is evidenced by Table 15, the final conclusions are basically of a qualitative nature.

A study of Appendix C reveals that the interpretation of the decision-trees is not always unambiguous. To reflect these uncertainties, several of the conclusions are qualified with "probably" or "possibly".

6.3 VALIDATION OF THE RESULTS FOR TEXAS LOCATIONS

A recent Texas Air Control Board (TACB) report (Price et al 1977) includes detailed analyses of the causes of high TSP levels at all 7 Texas locations considered in this study. The TACB report consists of a comprehensive analysis of several data bases including

- TSP measurements
- Trace element measurements from X-ray fluourescence
- Microscopy analyses
- Soil sample analyses
- Surface meteorological data
- Detailed site inspections
- Construction histories
- Emission inventories
- Traffic data

For most locations, the TACB report provides a rather convincing picture of the sources that contribute to high TSP levels.

Table 16 presents a summary of the TACB conclusions for the 7 Texas sites contained among our study locations. It is obvious that the TACB results include much more detail than the three types of sources (wind-blown dust, man-made dust, and conventional man-made sources) that we have attempted to identify by analyzing the decision-trees. It is also obvious that

---

* The reader is referred back to Figure 19 for the original decision-tree for Forrest City.

TABLE 14. INTERPRETATION OF DECISION-TREE FOR FORREST CITY, AR

SITE/CONCLUSION	CLASS	TERMINAL NODE #'s	AVERAGE TSP LEVEL ($\mu g/m^3$)	NUMBER OF DATA POINTS	QUALITATIVE DESCRIPTION
	All data	All 10	99.1	166	- - - - - - - - - - - - -
Forrest City, Arkansas Man-made sources (at least partly dust related) to the southwest or west. Also a possibility of contributions from wind-blown dust	I	1,2,3	156.1	34	high afternoon mixing height; very low wind speed or high afternoon mixing height; moderate to high wind speed; westerly wind; moderately dry ground; very low morning mixing height or high afternoon mixing height; moderate to high wind speed; southwesterly wind; moderately dry ground; moderate to high morning mixing height; warm night
	II	4	116.0	7	low afternoon mixing height; low morning mixing height
	III	5,6	102.3	29	high afternoon mixing height; moderate to high wind speed; southwesterly wind; moderately dry ground; very low morning mixing height; warm night or high afternoon mixing height; moderate to high wind speed; westerly wind; moderately dry ground; moderate to high morning mixing height; cool night
	IV	7,9	85.1	45	high afternoon mixing height; moderate to high wind speed; easterly wind or high afternoon mixing height; moderate to high wind speed; westerly wind; very wet ground
	V	8,10	69.2	51	low afternoon mixing height; moderate to high morning mixing height

TABLE 15. SUMMARY OF CONCLUSIONS FROM THE INTERPRETATION OF THE AID DECISION TREES

LOCATION	CONCLUSION
Forrest City, AR	Man-made sources (at least partly dust related) to the southwest or west. Also a possibility of contributions from wind-blown dust.
Helena, AR	A dominant man-made source to the southwest, probably dust related. Also, a possibility of contributions from wind-blown dust.
Jonesboro, AR	Man-made sources, some probably dust related. Also some minor contributions from wind-blown dust.
Springdale, AR	Man-made sources, mostly dust related. Also, some wind-blown dust. Dust sources may be most intense during Jan. to Aug. (possibly agricultural) and in the westerly direction.
Stuttgart (#1), AR	Dominant man-made source toward the north or northeast, mostly dust related, probably a point source. Also some minor contributions from wind-blown dust.
Stuttgart (#2), AR	Man-made dust sources. Also, some contributions from wind-blown dust.
Albuquerque (#1), NM	Man-made sources (mostly dust related). Also, a major contribution from wind-blown dust.
Albuquerque (#2), NM	Man-made sources, mostly dust related. Also wind-blown dust. Wind-blown dust sources and/or man-made dust sources may be concentrated toward the southeast.
Dona Ana, NM	Man-made dust sources, possibly concentrated toward the northeast. Possibly some minor contributions from wind-blown dust.
Las Cruces, NM	Man-made sources. Also, significant contributions from wind-blown dust. Dust sources may be most intense during winter and spring (possibly agriculture related).
Raton, NM	Man-made sources, mostly to the east. Also a possibility of wind-blown dust. Dust sources may be most intense in spring and least intense in winter (possibly agriculture related).
Bethany, OK	Major contributions from wind-blown dust. Also, some dust related man-made sources.
Muskogee, OK	Man-made dust sources, most intense from January to August (possibly agriculture related). Also a possibility of contributions from wind-blown dust.

TABLE 15. SUMMARY OF CONCLUSIONS FROM THE INTERPRETATION OF THE AID DECISION TREES (Continued)

LOCATION	CONCLUSION
Oklahoma City, OK	Man-made sources, mostly dust related. Sources are most intense toward the north and during the first half of the year (possibly agriculture related). Also a possibility of wind-blown dust.
Roger Mills Co., OK	Man-made sources, mostly dust related. Also a possibility of major contributions from wind-blown dust.
Sequoyah Co., OK	Man-made sources, mostly dust related. Also some minor contributions from wind-blown dust.
Tulsa (#1), OK	Major contributions from wind-blown dust. Also some contributions from man-made dust sources.
Tulsa (#2), OK	Significant contributions from wind-blown dust. Also, man-made dust sources which may be most intense toward the south or west.
Corpus Christi (#1), TX	Man-made sources, some dust related. Possibly some minor contributions from wind-blown dust.
Corpus Christi (#2), TX	Probably a mixture of man-made sources and wind-blown dust. Sources tend to be strongest toward the north or east.
Corpus Christi (#3), TX	Wind-blown dust and man-made sources (both dust related and conventional).
Harlingen, TX	Dust sources, both wind-blown and man-made.
Lubbock, TX	Substantial contributions from wind-blown dust. Significant dust sources (wind-blown or man-made) toward the west.
McAllen, TX	Three source types apparent: 1) general man-made dust sources, 2) wind-blown dust, and 3) man-made dust source to the north-northwest. Sources may be intensified during first half of year (possibly agriculture related).
San Benito, TX	Man-made sources, probably dust related and most prevalent during warm months (possibly agriculture related). Also, a possibility of contributions from wind-blown dust.

Table 16 and our results are not directly comparable because, for example, our definition of wind-blown dust would include part of the contributions from regional background, agricultural activity, construction activity, and several other TACB categories as well as all of the contribution from the TACB dust-storm category.

TABLE 16. SUMMARY OF TACB CONCLUSIONS CONCERNING THE 7 TEXAS LOCATIONS

	LOCATION*						
SOURCE	CC #1	CC #2	CC #3	HA	LU	MC	SB
Point Source Industrial Emissions		M					S
Fugitive Industrial Emissions		M					S
Fugitive Material-Handling Emissions		P					S
Unpaved Work Areas	M	M			M		
Unpaved Roads and Driveways	P	M	M	P	M	S	S
Traffic on Paved Streets	S		M		M	S	
Construction Activity			M	S	S		
Agricultural Activity				M	S	M	M
Dust Storms					S		M
Regional Background	S	S	S	S	S	S	S

P = Predominant (>50%) contribution to measured TSP

S = Significant (20%-50%) contribution to measured TSP

M = Minor (5%-20%) contribution to measured TSP

= Insignificant contribution to measured TSP

---

* See Table 1 for sites corresponding to these abbreviations

In order to provide a test of our conclusions, we have prepared, from the discussions in the TACB report, qualitative summaries that have a format similar to our conclusions. Table 17 compares these summaries of the TACB results to our conclusions. Although it is again evident that the TACB report is able to provide more detailed and definitive conclusions, the agreement between the TACB study and our study concerning the types of sources affecting each site is generally fair to good.

6.4 EVALUATION OF THE METHOD

This and the previous chapter have demonstrated that the AID decision-tree program is a very effective tool for characterizing the relationship between TSP concentrations and meteorology. It has also been demonstrated that the decision-trees can be interpreted to yield qualitative conclusions concerning the types of sources causing high TSP levels. Furthermore, for the 7 sites in Texas, we have found fairly good agreement, qualitatively, between our conclusions and the results of very detailed studies conducted for those sites by the TACB.

Despite this reasonable level of success, we do not recommend that the above method be used alone to classify sites as to cause of nonattainment for the purpose of regulatory planning. The reasons for this negative recommendation are as follows:

- The results of interpreting the AID decision-trees (ie. the TSP-meteorology relationship) are very qualitative. In some cases, the results are somewhat ambiguous in the sense that contributions from both wind-blown dust and man-made sources are apparent but there is little indication as to which predominates. There does not appear to be a reliable and accurate way of translating the results into quantitative statements concerning which sites would or would not exceed the TSP standards if wind-blown dust were discounted.

- Although the results of this study agree fairly well with the results of the comprehensive TACB study, the concurrence is by no means 100%. Greater reliability would be wanted if the results were to be used to support regulatory policy.

- The method used here is essentially directed at distinguishing wind-blown dust from man-made sources. A finer resolution of these

TABLE 17. COMPARISON OF TACB CONCLUSIONS TO THE CONCLUSIONS OBTAINED BY ANALYZING THE TSP/METEOROLOGY DECISION-TREES

LOCATION	CONCLUSION FROM DECISION-TREE	TACB CONCLUSION	AGREEMENT
Corpus Christi (#1)	Man-made sources, some dust-related. Possibly, some minor contributions from wind-blown dust.	Major contributions from man-made dust sources, particularly unpaved parking lots. Possibly, some very minor contributions from wind-blown dust.	Good
Corpus Christi (#2)	Probably a mixture of man-made sources and wind-blown dust. Sources tend to be strongest toward the north or east.	Major contributions from man-made sources, mostly dust-related but some conventional. Sources include storage piles, ship unloading, and industrial facilities. Some minor contributions from wind-blown dust.	Fair
Corpus Christi (#3)	Wind-blown dust and man-made sources (both dust related and conventional).	Major contributions from a variety of dust related sources, mostly man-made but some wind-blown.	Good
Harlingen	Dust sources, both wind-blown and man-made.	Major contributions from man-made dust sources, basically unpaved roads (towards the south) and construction activity. Some minor contributions from wind-blown dust.	Fair
Lubbock	Substantial contributions from wind-blown dust. Significant dust sources (wind-blown or man-made) toward the west.	Major contributions from wind-blown dust. Also some contributions from man-made dust sources. Much of the dust is likely related to agricultural and construction activities. Sources are most significant to the west.	Excellent

TABLE 17. COMPARISON OF TACB CONCLUSIONS TO THE CONCLUSIONS DETERMINED BY ANALYZING THE TSP/METEOROLOGY DECISION-TREES (Continued)

LOCATION	CONCLUSION FROM DECISION-TREE	TACB CONCLUSION	AGREEMENT
McAllen	Three source types apparent: (1) general man-made dust sources; (2) wind-blown dust; (3) man-made dust source toward the north-northwest. Sources may be intensified during the first half of the year (possibly agriculture related).	Major contributions from man-made dust sources, mostly fugitive dust from traffic sources. Also, some minor contributions from wind-blown dust.	Fair
San Benito	Man-made sources, probably dust related and most prevalent during warm months (possibly agriculture related). Also, a possibility of contributions from wind-blown dust.	Major contributions from man-made sources, mostly fugitive dust sources but some conventional sources. (The most significant sources are a cotton gin and grain storage facility that only operate during the summer. The sources are most significant to the south and southeast, with a secondary concentration of sources to the northeast.) Also, some significant contributions from wind-blown dust.	Good

> source types would be desirable for planning purposes. In particular, it would be desirable to know what part of wind-blown dust represents regional dust storms (which can reasonably be considered uncontrollable) and what part represents more localized wind-blown dust from agricultural fields, storage piles, construction sites, etc. (which may be amenable to some control).

Basically, although the relationship between TSP and meteorology provides useful insights as to causes of nonattainment, we do not think that it can stand alone in support of regulatory decisions. Classification of sites as to cause of nonattainment should be based on more information such as the additional factors considered in the TACB study. The only exceptions might be those locations, such as Lubbock TX, where the TSP-meteorology relationship points almost exclusively toward wind-blown dust.

7.0 REGIONAL PATTERNS IN THE CONTRIBUTION OF WIND-BLOWN DUST TO TSP CONCENTRATIONS

This chapter provides an overall view of the contribution of wind-blown dust to high TSP concentrations in EPA Region 6 . The patterns in wind-blown dust contributions are examined with respect to both geographical features and site characteristics. Before examining these patterns, however, an alternative method of quantifying the TSP-meteorology relationship is applied. The results of this alternative method are combined with the results of the AID analysis to yield our best estimates of wind-blown dust contributions at the study locations.

7.1 AN ALTERNATIVE ANALYSIS OF THE TSP-METEOROLOGY RELATIONSHIP

The previous two chapters conducted a detailed study of the TSP-meteorology relationship and drew inferences concerning the types of sources contributing to high TSP concentrations. Recognizing that wind speed and precipitation are two of the key parameters for identifying wind-blown dust, this section performs a simpler analysis of the data in an attempt to identify days of high TSP caused by wind-blown dust. The new method is based on a two-dimensional nomogram involving wind speed and precipitation.

Table 18 presents the nomogram that is intended to classify days of high TSP as to the probability of the source being wind-blown dust. The ranges in the nomogram have been selected based on an examination of the raw data (including days when wind-blown dust was actually observed), on discussions with the EPA Region 6 Meteorologist (Hall 1978), and on insight based on experience. The advantages of using the nomogram to classify high TSP days as to probability of wind-blown dust contributions are that it is simple, has minimal data requirements, and yields quantitative results. The disadvantage, of course, is that it may be inaccurate. More variables (eg. mixing height, relative humidity, other precipitation parameters, etc.) may actually be required to identify wind-blown dust situations with good precision, and the ranges of the three variables that we have chosen may not be specified correctly. In particular, we note that the wind and dryness conditions required for wind-blown dust are not absolute; rather they can be very sensitive to the condition of the top soil.

TABLE 18. NOMOGRAM FOR CLASSIFYING DAYS OF HIGH TSP ACCORDING TO PROBABILITY OF WIND-BLOWN DUST

	Precipitation During Prior Two Days		
Maximum and Average Wind Speeds	PREC23≤.01	.01≤PREC23<.50	.50<PREC23*
WSMAX≥35 and AVWS≥13	A	A	B
WSMAX<35 or AVWS<13 and WSMAX≥25 and AVWS≥10	A	B	C
WSMAX<25 or AVWS<10 and WSMAX>16 or AVWS>7.5	B	C	D
WSMAX≤16 and AVWS≤7.5 and WSMAX≥12 or AVWS>6	C	D	E
WSMAX≤12 and AVWS≤6	D	E	E

A = Wind-blown dust

B = Probably wind-blown dust

C = Inconclusive

D = Probably not wind-blown dust

E = Not wind-blown dust

*PREC23 = PREC3 - PREC1

Table 19 presents the distributions of all sampling days into the classes defined by the nomogram. As indicated by the averages at the bottom of Table 19, the distribution of all days is slightly skewed toward the left (windier and drier classes). Considering the definitions of the classes (Table 18), this reflects the very windy and very dry conditions that tend

TABLE 19. DISTRIBUTION OF ALL SAMPLING DAYS INTO FIVE CLASSES FOR DISTINGUISHING PROBABILITY OF WIND-BLOWN DUST

LOCATION	Total Number of Samples	Distribution of Samples by Class				
		A	B	C	D	E
Forrest City, AR	166	4	46	57	40	19
Helena, AR	168	4	47	56	43	18
Jonesboro, AR	165	2	48	54	41	20
Springdale, AR	143	0	26	37	52	28
Stuttgart (#1), AR	168	4	47	56	42	19
Stuttgart (#2), AR	126	4	32	43	33	14
Albuquerque (#1), NM	96	11	39	24	18	4
Albuquerque (#2), NM	95	12	33	24	21	5
Dona Ana, NM	135	9	47	46	30	3
Las Cruces, NM	113	8	38	35	28	4
Raton, NM	85	43	29	9	4	0
Bethany, OK	186	22	73	48	30	13
Muskogee, OK	217	19	90	45	41	22
Oklahoma City, OK	137	18	58	32	22	7
Roger Mills Co., OK	106	15	38	23	19	11
Sequoyah, OK	119	0	19	33	39	28
Tulsa (#1), OK	210	16	83	43	46	22
Tulsa (#2), OK	232	19	89	47	51	26
Corpus Christi (#1), TX	128	22	60	29	9	8
Corpus Christi (#2), TX	121	25	54	33	5	4
Corpus Christi (#3), TX	106	18	44	29	7	8
Harlingen, TX	149	13	69	35	20	7
Lubbock, TX	108	18	51	20	14	5
McAllen, TX	125	15	55	28	17	10
San Benito, TX	136	15	60	29	23	9
AVERAGE OF ALL SITES	141.6	13.6	51.0	36.6	27.8	12.6

to prevail in most of the study area. The distributions are more heavily skewed to the left for the sites in New Mexico, western Oklahoma, and southern or western Texas and either centered or skewed slightly to the right for the sites in Arkansas and eastern Oklahoma. This spatial pattern, of course, just reflects the geographical distribution of wind speed and rainfall (see Figures 5, 6, and 9 in Chapter 2).

Table 20 presents the distribution of days with high TSP among the five classes. There is considerable variety in the results among the study locations. Many of the sites have very skewed distributions (right or left) for the high TSP days, indicating that high TSP levels at those locations are either dominated by wind-blown dust or have little to do with wind-blown dust.

Table 21 summarizes the results of the AID and nomogram analyses by ranking each site, in a semi-quantitative manner, according to the contribution of wind-blown dust to high TSP concentrations. The ranking is based on an index from 0 to 4 representing a range of conclusions from "high TSP levels basically all not from wind-blown dust" to "high TSP levels basically all from wind-blown dust". Ranges for the rankings are often given in Table 21 to reflect the uncertainties in the conclusions from the AID and nomogram results.

As evidenced by Table 21, the results of the nomogram analysis are often considerably different from the results of the AID analysis. The basic area of disagreement is that the nomogram analysis tends to implicate wind-blown dust sources much more strongly than does the AID analysis.

There are two basic explanations for the disagreement. First, the conclusions of both the nomogram and AID analyses could be correct, with each applying to different definitions of what constitutes "high TSP levels". The nomogram analysis is based only on days when TSP concentrations exceed the primary or secondary air quality standards (260 $\mu g/m^3$ and 150 $\mu g/m^3$, respectively). Although we tended to focus greatest attention on the terminal nodes with highest TSP concentrations in interpreting the AID decision trees, the AID analysis is essentially based on all the TSP data. It is quite possible, and actually physically reasonable, that the few days of extremely high TSP levels are more dominated by wind-blown dust than are the

TABLE 20. DISTRIBUTION OF HIGH TSP LEVELS AS TO PROBABILITY OF WIND-BLOWN DUST ACCORDING TO THE NOMOGRAM

LOCATION	All Data	Wind-Blown Dust	Probably Wind-Blown Dust	Inconclusive	Probably Not Wind-Blown Dust	Not Wind-Blown Dust
Forrest City, AR	2* 21**	0 1	2 11	0 2	0 5	0 2
Helena, AR	1 19	0 1	0 8	1 7	0 1	0 2
Jonesboro, AR	0 6	0 0	0 0	0 1	0 3	0 2
Springdale, AR	0 3	0 0	0 0	0 1	0 0	0 2
Stuttgart (#1), AR	2 24	0 0	1 2	0 6	0 11	1 5
Stuttgart (#2), AR	25 75	1 3	7 21	8 27	8 21	1 3
Albuquerque (#1), NM	1 21	0 4	0 6	0 5	1 6	0 0
Albuquerque (#2), NM	0 8	0 0	0 3	0 4	0 1	0 0
Dona Ana, NM	2 21	2 2	0 9	0 5	0 5	0 0
Las Cruces, NM	2 7	0 1	0 1	1 3	1 2	0 0
Raton, NM	0 1	0 0	0 1	0 0	0 0	0 0
Bethany, OK	0 12	0 6	0 5	0 0	0 1	0 0

* # Days >260 $\mu g/m^3$

** # Days >150 $\mu g/m^3$

TABLE 20. DISTRIBUTION OF HIGH TSP LEVELS AS TO PROBABILITY OF WIND-BLOWN DUST ACCORDING TO THE NOMOGRAM (Cont'd.)

LOCATION	All Data	Wind-Blown Dust	Probably Wind-Blown Dust	Inconclusive	Probably Not Wind-Blown Dust	Not Wind-Blown Dust
Muskogee, OK	0* 3**	0 1	0 2	0 0	0 0	0 0
Oklahoma City, OK	0 3	0 3	0 0	0 0	0 0	0 0
Roger Mills Co., OK	3 8	0 2	2 4	0 1	1 1	0 0
Sequoyah, OK	0 3	0 0	0 0	0 0	0 3	0 0
Tulsa (#1), OK	5 16	3 5	2 6	0 3	0 2	0 0
Tulsa (#2), OK	1 6	1 2	0 1	0 1	0 2	0 0
Corpus Christi (#1), TX	1 3	1 1	0 2	0 0	0 0	0 0
Corpus Christi (#2), TX	3 24	0 8	0 10	3 5	0 1	0 0
Corpus Christi (#3), TX	0 1	0 1	0 0	0 0	0 0	0 0
Harlingen, TX	3 19	2 5	1 13	0 1	0 0	0 0
Lubbock, TX	5 12	3 6	2 5	0 0	0 1	0 0
McAllen, TX	1 15	0 3	1 6	0 2	0 4	0 0
San Benito, TX	14 32	2 5	7 15	1 5	1 4	3 3

* # Days > 260 μg/m^3
** # Days > 150 μg/m

TABLE 21. COMPARISON OF THE RESULTS FROM THE AID AND NOMOGRAM ANALYSES CONCERNING CONTRIBUTIONS OF WIND-BLOWN DUST TO TSP

Location	Ranking Based on AID Analysis	Ranking Based On Nomogram Analysis
Forrest City, AR	0 - 2	2 - 3
Helena, AR	0 - 2	2 - 3
Jonesboro, AR	1	1
Springdale, AR	1 - 2	0 - 2
Stuttgart (#1), AR	1	1 - 2
Stuttgart (#2), AR	1 - 2	2
Albuquerque (#1), NM	3	2 - 3
Albuquerque (#2), NM	2 - 3	2 - 3
Dona Ana, NM	0 - 1	2 - 3
Las Cruces, NM	2	2
Raton, NM	0 - 2	1 - 4
Bethany, OK	3	3 - 4
Muskogee, OK	0 - 2	3 - 4
Oklahoma City, OK	0 - 2	4
Roger Mills Co., OK	1 - 3	3
Sequoyah, OK	1	1
Tulsa (#1), OK	3	3
Tulsa (#2), OK	3	2 - 3
Corpus Christi (#1), TX	0 - 1	3 - 4
Corpus Christi (#2), TX	1 - 3	3
Corpus Christi (#3), TX	2	2 - 4
Harlingen, TX	2	3
Lubbock, TX	3 - 4	3 - 4
McAllen, TX	1 - 2	2 - 3
San Benito, TX	0 - 2	2 - 3

4 = High TSP levels basically all from wind-blown dust
3 = High TSP levels mostly from wind-blown dust
2 = High TSP levels approximately half from wind-blown dust
1 = High TSP levels mostly not from wind-blown dust
0 = High TSP levels basically all not from wind-blown dust

more normal variations in TSP levels. If this is the case, we would conclude that wind-blown dust is significantly more important in relation to high short-term (e.g. daily) TSP concentrations than it is in relation to high long-term (e.g. yearly) TSP concentrations.

The second explanation for the disagreement is that the nomogram results and/or the AID results are partly in error. We would tend to place greater confidence in the AID results because the AID analysis uses all the sampling days, not just a few, and because the AID analysis considers all 18 meteorological variables, not just three. In fact, the discrepancies may imply that a 2-dimensional nomogram is too simplistic for discerning wind-blown dust situations with good precision. However, the nomogram is much easier to interpret than are the AID decision trees, and, if the errors lie in interpretations, then the AID analysis might be the one at fault.

Resolving the causes of the disagreements would require considerably more work than is possible within the level of effort allocated to this study. So that we can continue with the analysis and examine patterns in the contribution of wind-blown dust to TSP, we have decided to formulate a simple 3-class categorization of the sites:

I. Sites where wind-blown dust appears to be only a minor contributor to high TSP levels.

II. Sites where wind-blown dust appears to be a significant but not the major contributor to high TSP levels, or sites where there is large uncertainty.

III. Sites where wind-blown dust appears to be the major contributor to high TSP levels.

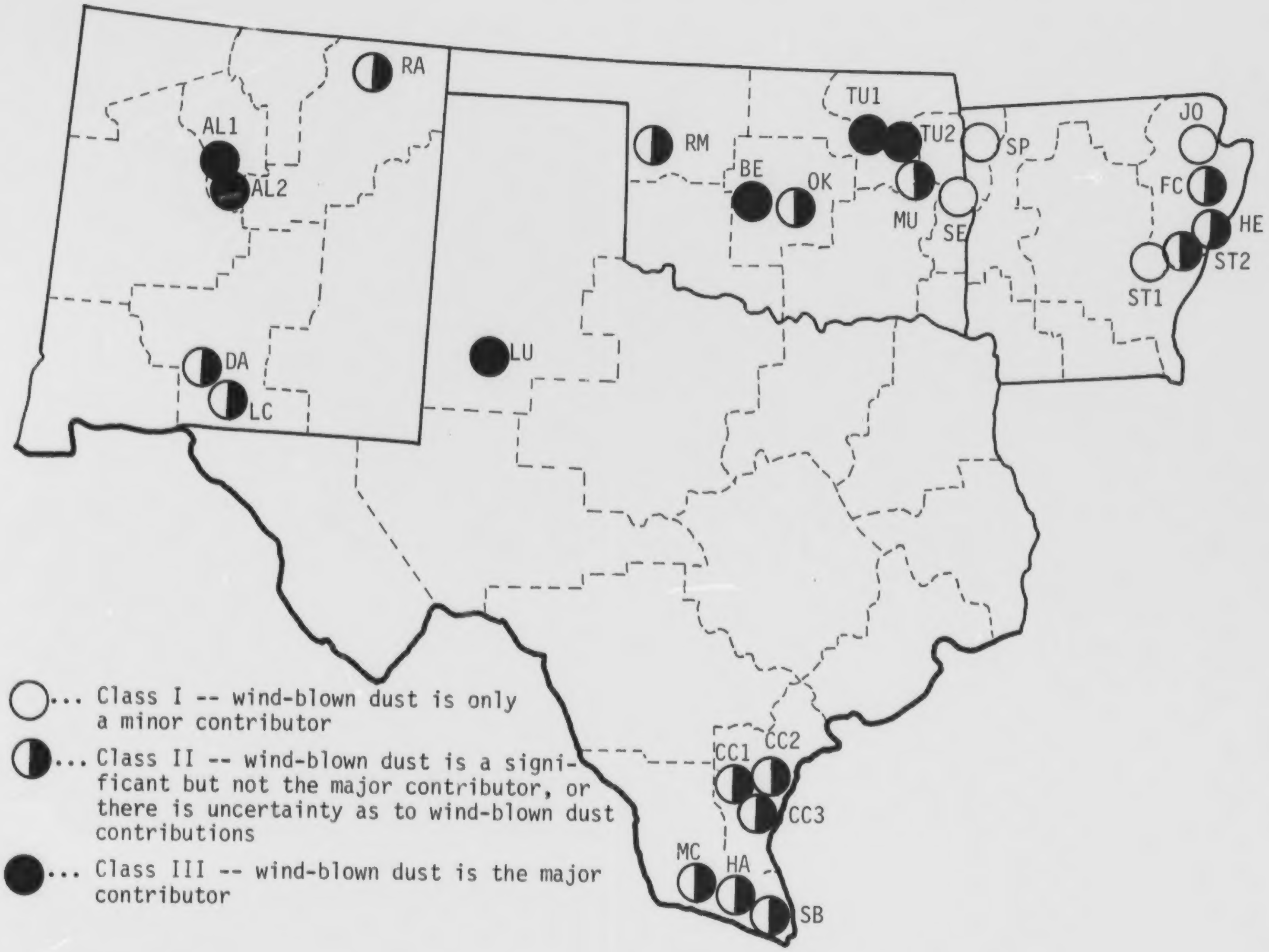
This categorization, presented in Table 22, is based on both the AID and the nomogram results, with consideration given to the uncertainties within each analysis and to the disagreements or agreements between the analyses. Partly because of the uncertainties, most sites fall into Class II.

7.2 PATTERNS IN CAUSES OF HIGH TSP LEVELS

Figure 20 illustrates the geographical distribution of the sites classified according to causes of high TSP levels. Three of the four

TABLE 22. THREE-CLASS CATEGORIZATION OF SITES WITH RESPECT TO WIND-BLOWN DUST CONTRIBUTION.

LOCATION	CATEGORIZATION
Forrest City, AR	II
Helena, AR	II
Jonesboro, AR	I
Springdale, AR	I
Stuttgart (#1), AR	I
Stuttgart (#2), AR	II
Albuquerque (#1), NM	III
Albuquerque (#2), NM	III
Dona Ana, NM	II
Las Cruces, NM	II
Raton, NM	II
Bethany, OK	III
Muskogee, OK	II
Oklahoma City, OK	II
Roger Mills Co., OK	II
Sequoyah, OK	I
Tulsa (#1), OK	III
Tulsa (#2), OK	III
Corpus Christi (#1), TX	II
Corpus Christi (#2), TX	II
Corpus Christi (#3), TX	II
Harlingen, TX	II
Lubbock, TX	III
McAllen, TX	II
San Benito, TX	II



RA
AL1
AL2
DA
LC
LU
RM
BE
OK
TU1
TU2
MU
SE
SP
JO
FC
HE
ST2
ST1
CC1
CC2
CC3
MC
HA
SB
○... Class I -- wind-blown dust is only a minor contributor
◐... Class II -- wind-blown dust is a significant but not the major contributor, or there is uncertainty as to wind-blown dust contributions
●... Class III -- wind-blown dust is the major contributor

Figure 20. Geographical distribution of sites classified according to causes of high TSP levels.

locations in Class I (not wind-blown dust) are in Arkansas, and the other is in Oklahoma near the Arkansas border. The six locations in Class III (wind-blown dust) occur in a band from central Oklahoma, through northern Texas, to central New Mexico. Sites in Class II (indeterminant) are scattered over the entire study region.

Figure 20 appears to be fairly consistent with the results of Hagen and Woodruff (1973) and Orgill and Sehmel (1976). Hagen and Woodruff used airport data on blowing dust observations, visibility, and wind speed to estimate the frequency of wind-blown dust (see Figure 21). Orgill and Sehmel used only the recorded observations of blowing-dust (see Figure 22). As shown in Figures 21 and 22, the Texas panhandle is the "hot-spot" for wind-blown dust in Region 6, with significant occurrences also apparent in central/eastern New Mexico, western Oklahoma, and the remainder of western Texas. Wind-blown dust is relatively rare in Arkansas and Louisiana.

From the geographical patterns in Figure 20, we venture the following gross generalizations: At Region 6 locations in the Texas Panhandle or within approximately 150 miles of the Texas Panhandle, wind-blown dust tends to be a significant contributor, and often the dominant contributor, to high TSP concentrations. In Arkansas (and, based on climatology, probably Louisiana), wind-blown dust tends to be a minor, often insignificant, contributor to high TSP concentrations. At other Region 6 locations, wind-blown dust may or may not be a significant source of high TSP levels.

In an attempt to gain further insight as to the types of sites that have major or minor contributions from wind-blown dust, the categorization of the sites (into Classes I, II, and III) was compared to the list of site characteristics summarized in Table 2 (page14). We found _no_ obvious correlations between contributions from wind-blown dust and SAROAD site type (e.g. center-city versus suburban versus rural, or commercial versus industrial versus residential versus agricultural), county wide emission density, or elevation of monitor above ground. The only pattern that appears to exist involves climatology; specifically, wind-blown dust is positively correlated with high wind speed and low precipitation conditions. The correlation with climatology is, of course, closely related to the geographical patterns in wind-blown dust contributions.

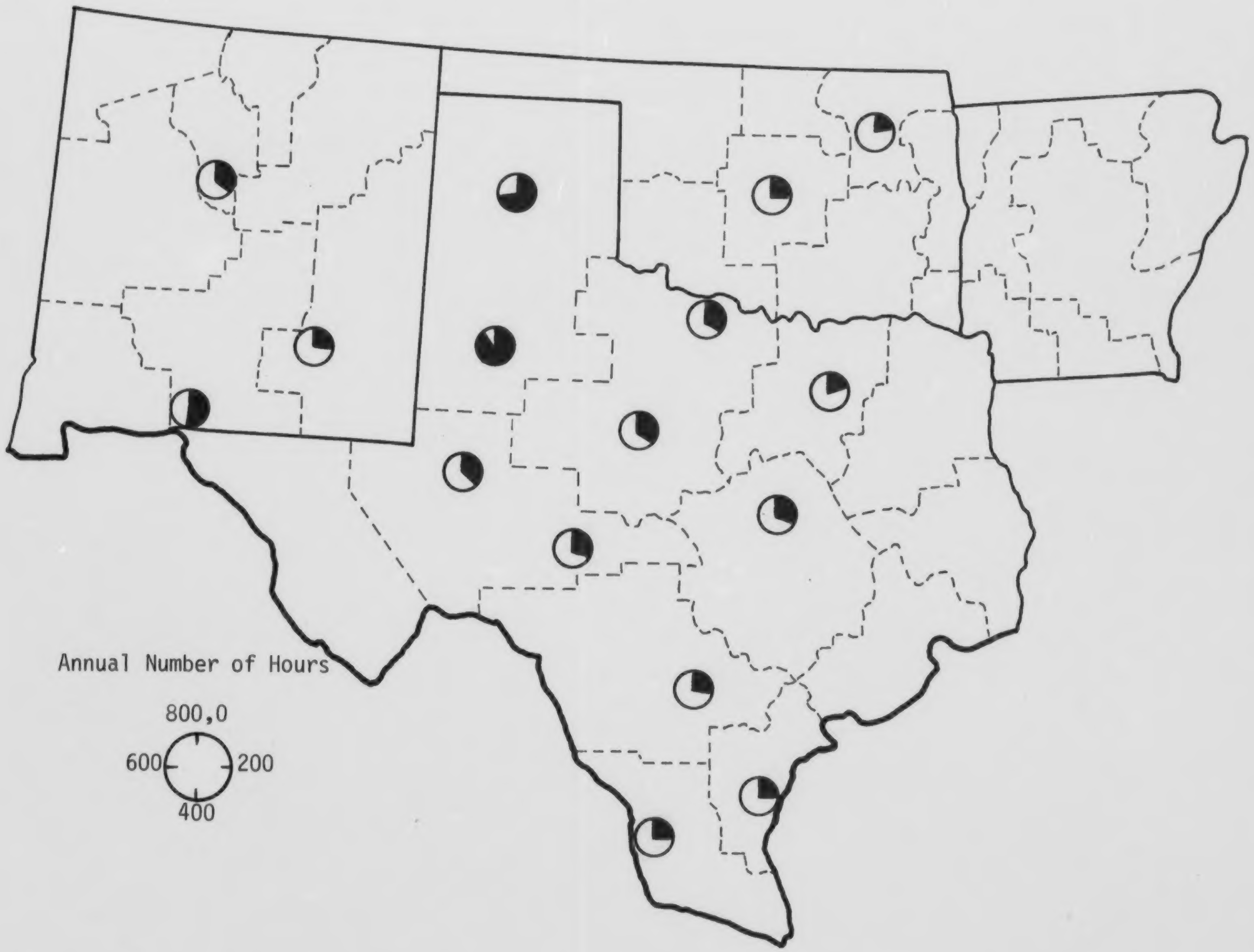

Annual Number of Hours
800,0
200
400
600

Figure 21. Annual average number of hours with wind-blown dust during the 1950's (Hagen and Woodruff 1973).

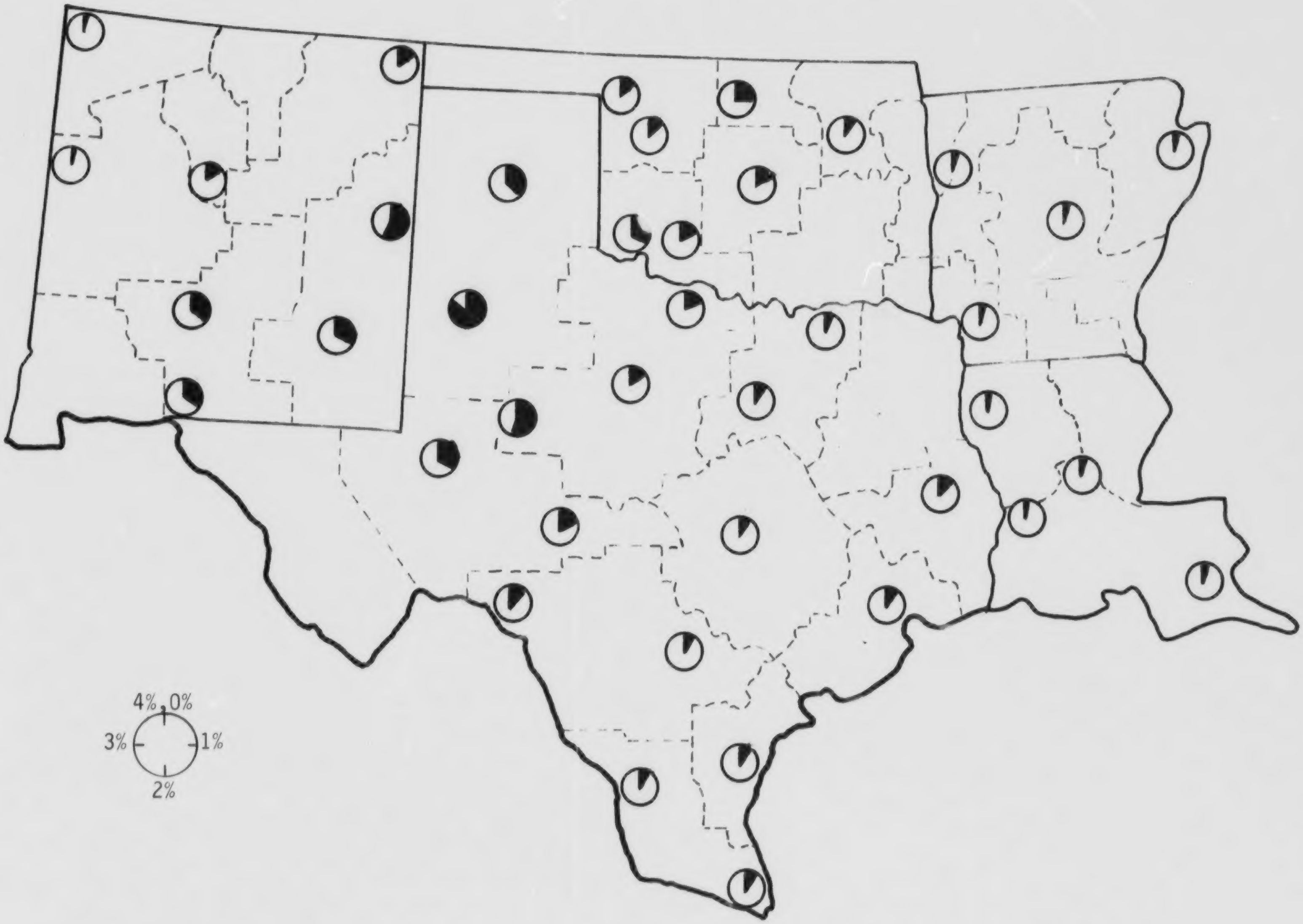


Figure 22. Annual percent frequency of hours with blowing dust observations, 1940-1970 (Orgill and Sehmel 1976).

8.0 TIME SERIES OF TSP DATA

As part of air quality planning for TSP, it is important to analyze the ambient data as a time series. Analyses of trends over time can indicate whether control strategies are improving air quality or whether source growth is worsening air quality. Periods of especially high TSP levels may reflect temporary source activity such as construction. This chapter briefly examines the time series for TSP at the 25 study locations from 1973 to 1975 and shows how the AID decision-tree results can be useful in characterizing historical trends in TSP concentrations.

8.1 TIME SERIES PLOTS AND REGRESSION ANALYSIS

Appendix D presents time series plots of the TSP data for 1973-1975 at the 25 study locations. As evidenced by a cursory examination of these plots, the large meteorological variance in daily TSP concentrations makes interpretation of the time series very difficult. Scrutiny of the plots, however, leads to the following observations: There may have been upward trends in TSP levels at Albuquerque (#2), Corpus Christi (#1), and Corpus Christi (#3), and downward trends at Forrest City, Helena, Springdale, Stuttgart (#2), Roger Mills, and Tulsa (#1). There may have been temporary source activity (e.g. construction) at Forrest City in 1973, Muskogee in 1974, Harlingen in 1974, and Lubbock in 1974-early 1975. The most obvious of all these time series patterns is the high level of TSP at Harlingen in 1974 (see Figure 23).

To gain further insight as to the possibility of significant overall trends during 1973-1975, regression analyses have been performed at each site with TSP as the dependent variable and time as the independent variable. A trend is interpreted as significant if the t-statistic of the regression coefficient exceeds 2.0 (approximately a 95% confidence level). The five sites showing significant upward trends in TSP concentrations are as follows (t-statistics in parentheses): Albuquerque #2 (4.5), Muskogee (2.0), Corpus Christi #1 (3.0), Corpus Christi #3 (2.8), and Harlingen (2.1). The four sites showing significant downward trends, all of which are

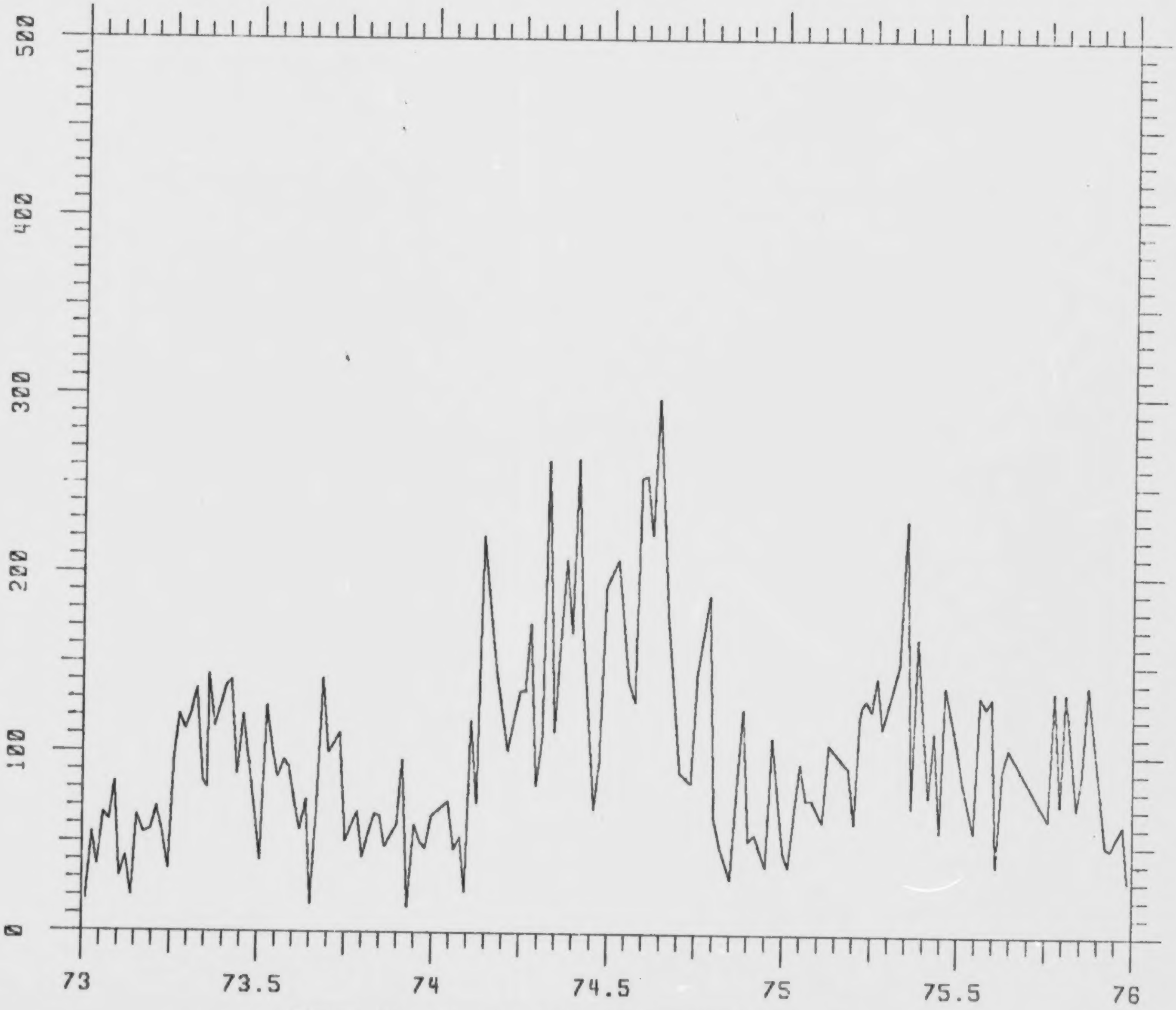


Figure 23. Time series plot for Harlingen, TX.

located in Arkansas, are Forrest City (4.1), Helena (2.4), Springdale (3.7), and Stuttgart #2 (2.8).

It should be noted that a statistically significant trend does not imply that the trend was produced by emission changes. Weather fluctuations can produce quite substantial changes in air quality, especially when dealing with only a three-year period, and especially in the case of a pollutant that is sampled infrequently (such as TSP). To determine the actual amount of air quality change that is due to emissions, the TSP trends should be normalized for weather changes. The next section discusses a method for performing this normalization.

8.2 METEOROLOGICAL ADJUSTMENT OF AMBIENT TSP DATA

Meteorological variance in daily TSP levels hinders the interpretation of the time series data. The results of the AID program provide a means of circumventing this problem by examining data within fixed meteorological classes. Stratifying the time series by meteorological class should eliminate some of the meteorological bias in the trends and should facilitate interpretation of the trends by reducing the variance in the data.

Figure 24 illustrates the Harlingen, TX time series stratified into four meteorological classes. These meteorological classes, defined in Appendix C, have been constructed by combining the 10 terminal nodes of the AID decision-tree into 4 groups. The meteorologically stratified data still show an apparent increase in TSP levels during 1974, during the months February through October. Accordingly, we conclude that there probably was a temporary TSP source, such as construction, during that period at Harlingen.

It is evident in Figure 24 (and in similar plots for the other locations) that significant variance exists within each meteorological class and that the values in the classes overlap considerably. The explanation is that the AID program typically explains only about 50% of the variance in TSP levels; also, some of this resolution is lost when the 10 terminal nodes are synthesized into 4 or 5 classes. One might consider using all

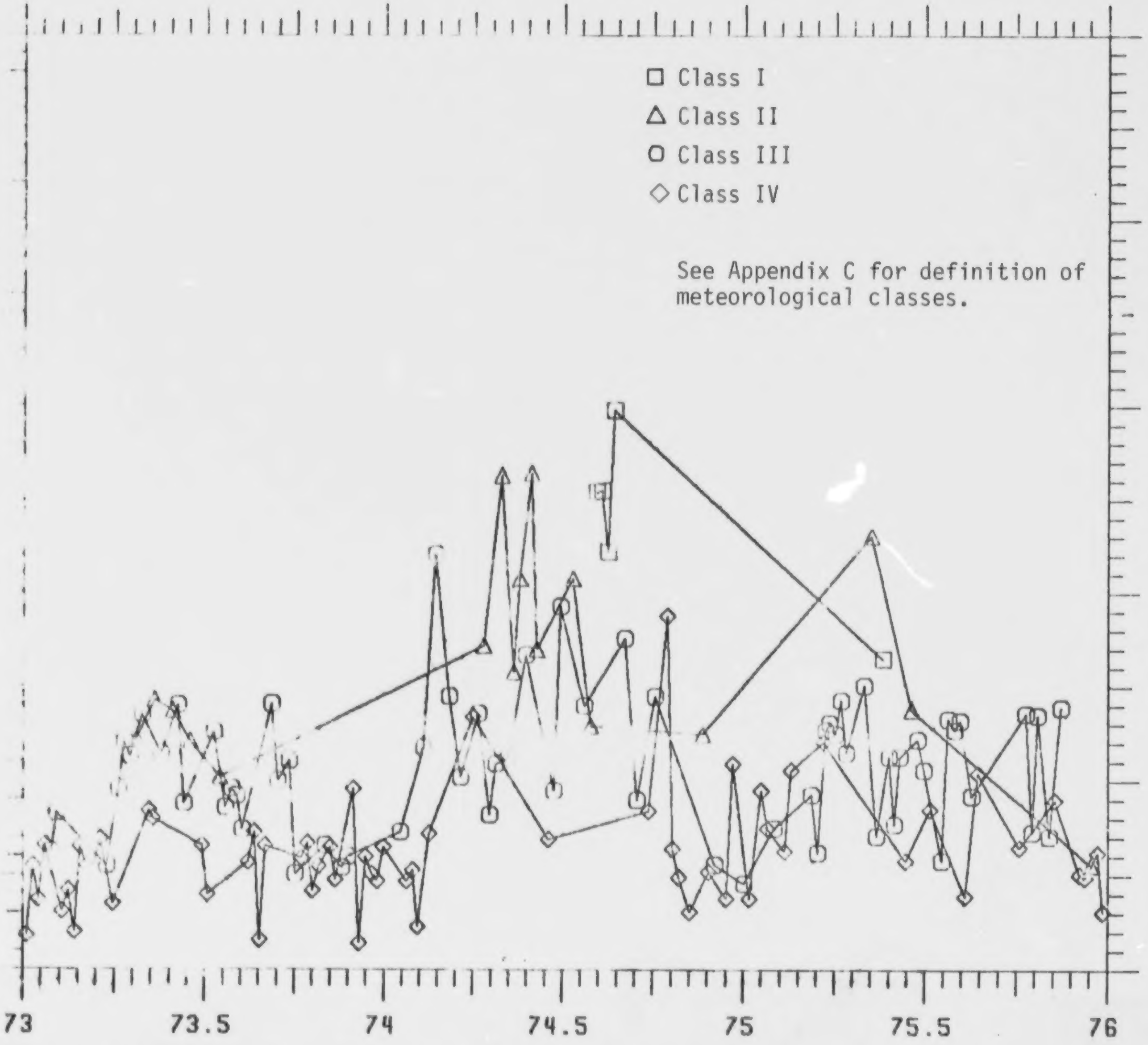


Figure 24. Time series at Harlingen TX, stratified by meteorological class.

10 terminal nodes as individual meteorological classes, but then one would be faced with the problem of even smaller sample sizes within each class.

The meteorological classes derived from the AID results are theoretically well suited for a quantitative meteorological normalization scheme proposed by Zeldin and Meisel (1977). Zeldin and Meisel showed that meteorological normalization of air quality statistics (such as the arithmetic mean, geometric mean, frequency of standard violations, etc.) can be performed in a straightforward manner if one is able to determine meteorological classes that explain a significant fraction of air quality variance. For example, a meteorologically normalized annual mean can be computed as follows:

$$\bar{Y}_{k,NORM} = \sum_{i=1}^{M} \bar{Y}_{ki} P_i$$

where $\bar{Y}_{k,\ NORM}$ = the meteorologically normalized mean for a pollutant Y in the kth year,

M = the number of meteorological classes,

$\bar{Y}_{ki}$ = the average value of Y for data in meteorological class "i" during the kth year,

and P_i = the average frequency of occurrence of meteorological class "i" during all years.

We have applied this meteorological normalization scheme to annual mean TSP levels at Harlingen, TX. As illustrated in Figure 25, the results indicate that the high TSP levels in 1974 were partly due to emissions (because the normalized data show a peak in 1974) and partly due to meteorology (because the 1974 peak in the normalized data is less than the 1974 peak in the actual data). These conclusions agree with the assessment of Price et al. (1977) who found that there was major construction activity near the site in August and September 1974 and that extremely dry weather conditions (conducive to dust) prevailed at Harlingen from February through July 1974.

As a caveat to the seemingly successful application of meteorological adjustment procedures at Harlingen, we note that the data bases in this study are really not suited for meteorological normalization analyses. Zeldin and Meisel have remarked that at least 5 years of data (possibly

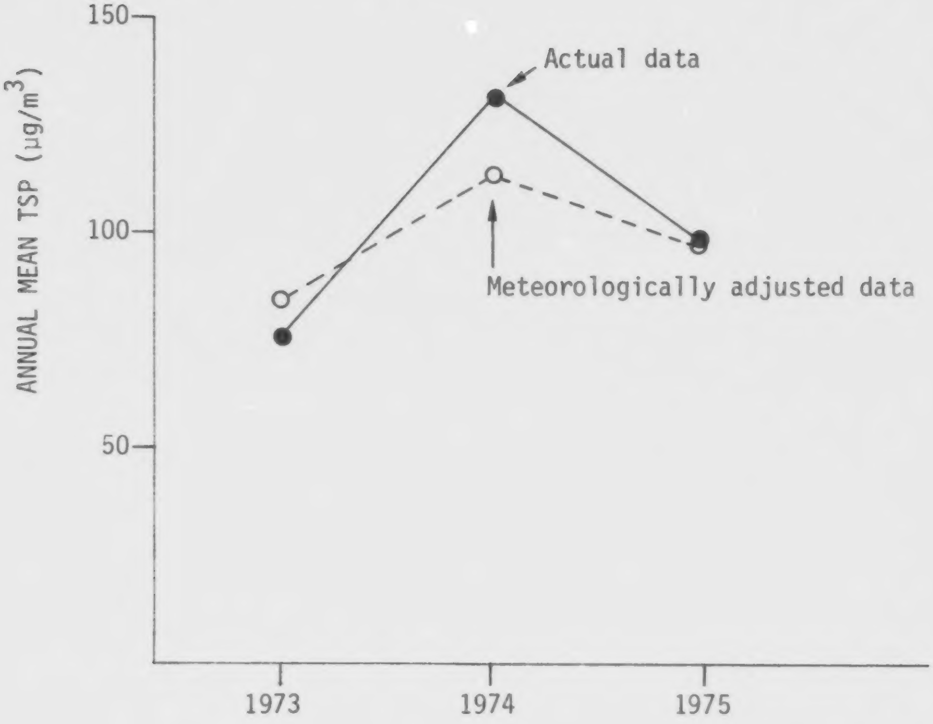


Figure 25. Meteorological adjustment of annual mean TSP concentrations at Harlingen, TX.

more for an intermittently sampled pollutant such as TSP) should be available in order to apply their meteorological adjustment scheme. Also, the meteorological classes should be derived with only a portion of the available data and checked with an independent data set. There is a very distinct possibility that the meteorological classes at Harlingen may have been biased by higher TSP emissions in 1974 (i.e. weather conditions that occurred most frequently in 1974 were selected as the conditions most conducive to high TSP). That this type of bias will likely occur in working with small data sets is supported by similar analyses that we conducted for other locations that had one exceptionally high year of TSP (we analyzed all the locations listed at the end of Chapter 4). For almost every site, the meteorological normalizations lowered the value for the high year slightly and increased the value for the low years slightly (but not enough so that the high and low years were equal). This disturbingly regular pattern in the results of the meteorological normalizations may be a product of biases introduced in the meteorological classes by changes in emission levels.

The above remarks lead to an even more general comment concerning time series analysis of TSP data. Because TSP is sampled infrequently, and because weather conditions lead to large variance in TSP concentrations, it seems advisable to have at least 5 to 10 years of TSP data before conducting detailed analyses of the time series. With larger data sets, much more confidence can be placed in trend analyses and in meteorological normalization studies.

9.0 REFERENCES

Curran, T.C. and N.H. Frank, "Assessing the Validity of the Lognormal Model when Predicting Maximum Air Pollution Concentrations," 68th Meeting of the Air Pollution Control Association, Boston, 1975

EPA, "Monitoring and Air Quality Trends Report, 1973," EPA - 450/1-74-007, Office of Air Quality Planning and Standards, 1974.

EPA, "Directory of Air Quality Monitoring Sites Active in 1974," EPA - 450/2-76-008, Office of Air Quality Planning and Standards, 1976.

ESSA (Environmental Science Services Administration of the U.S. Department of Commerce), "Climatic Atlas of the United States," Ashville, N.C., 1968.

Flores, M., EPA Region 6 , Dallas, Texas, Personal Commumication, Oct. 1977.

Gifford, F.A., "The Form of the Frequency Distribution of Air Pollution Concentrations," Proceedings of the Symposium on Statistical Aspects of of Air Quality Data, EPA Document 650/4-74-038, EPA Office of Research and Development, 1974.

Hall, F., Regional Meteorologist, EPA Region 6 , Dallas, Texas, Personal Communication, October 1977 and April 1978.

Hagen, L.J. and N.P. Woodruff, "Air Pollution from Dust Storms in the Great Plains," Atmospheric Environment, Volume 7, p. 323, 1973.

Larsen, R.I., C.E. Zimmer, D.A. Lynn, and K.G. Bemel, "Analyzing Air Pollution Concentration and Dosage Data," Journal of the Air Pollution Control Association, Vol. 17, p. 85, 1967.

Larsen, R.I.,"A New Mathematical Model of Air Pollutant Concentration Averaging Time and Frequency," Journal of the Air Pollution Control Association, Vol. 28, p. 24, 1969.

Larsen, R.I., A Mathematical Model for Relating Air Quality Measurements to Air Quality Standards, Publication AP-89, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, 1971.

Larsen, R.I., "An Air Quality Data Analysis System for Interrelating Effects, Standards, and Needed Source Reductions," Journal of the Air Pollution Control Association, Vos. 23, P. 933, 1973

Larsen, R.I., "An Air Quality Data Analysis System for Interrelating Effects, Standards, and Needed Source Reductions - Part 2," Journal of the Air Pollution Control Association, Vol. 24, P. 551, 1974.

Meisel, W.S. and T.E. Dushane, "Monitoring Carbon Monoxide in Urban Areas," Prepared for National Cooperative Highway Research Program by Technology Service Corporation, 1977.

NCC (National Climatic Center), "Local Climatological Data," Summaries for each month from 1973 to 1975, 1977.

Neustadter, H.E., S.M. Sidik, and J.C. Burr, Jr., "Statistical Summary and Trend Evaluation of Air Quality Data for Cleveland, Ohio in 1967 to 1971: Total Suspended Particulate, Nitrogen Dioxide, and Sulfur Dioxide," NASA TN D-6935, Lewis Research Center, Cleveland, Ohio, 1973.

Orgill, M.M. and G.A. Sehmel, "Frequency and Diurnal Variation of Dust Storms in the Contiguous U.S.A.," Atmospheric Environment, Volume 10, p. 813, 1976.

Pace, T., EPA Office of Air Quality Planning and Standards, Research Triangle Park, North Carolina, Personal Communication of data, October 1977.

Price, J.H., J.P. Gise, H.E. Sievers, S.E. Ehlers, and B.K. Knape, "Attainment Analysis, Volume I, Causes of Nonattainment," Texas Air Control Board, Austin, Texas, January 1977.

Record, F.A., et al, "National Assessment of the Urban Particulate Problem, Volume I: Summary of National Assessment," Prepared for EPA Office of Air Quality Planning and Standards by GCA/Technology Division, EPA - 450/3-76-024, 1976.

Richard, G., et al, "An Implementation Plan for Suspended Particulate Matter in the Phoenix Area," Vols. 1-4, Prepared for EPA Office of Air Quality Planning and Standards by TRW Environmental Engineering Under Contract Number 68-01-3152, 1977.

Singpurwalla, N.D., "Extreme Values from a Lognormal Law with Applications to Air Pollution Problems," Technometrics, Vol. 14, P. 703, 1972.

Sonquist, J.A., E.L. Baker, and J.N. Morgan, "Searching for Structure," Institute for Social Research, the University of Michigan, 1973.

Trijonis, J.C., "Empirical Studies of Ambient Nitrogen Dioxide Air Quality and NO_2/Precursor Relationships," Prepared for EPA Office of Research and Development by Technology Service Corporation Under Contract Number 68-02-2299, 1977.

Zeldin, M.D. and W.S. Meisel, "Guideline Document on Use of Meteorological Data in Air Quality Trend Analysis," Prepared for EPA Office of Air Quality Planning and Standards by Technology Service Corporation under Contract No. 68-02-2318, November 1977.

Zimmer, C.E., and R.I. Larsen, "Calculating Air Quality and Its Control," Journal of the Air Pollution Control Association, Vol. 15, p. 565, 1965.

APPENDIX A

CUMULATIVE FREQUENCY DISTRIBUTIONS OF TSP DATA

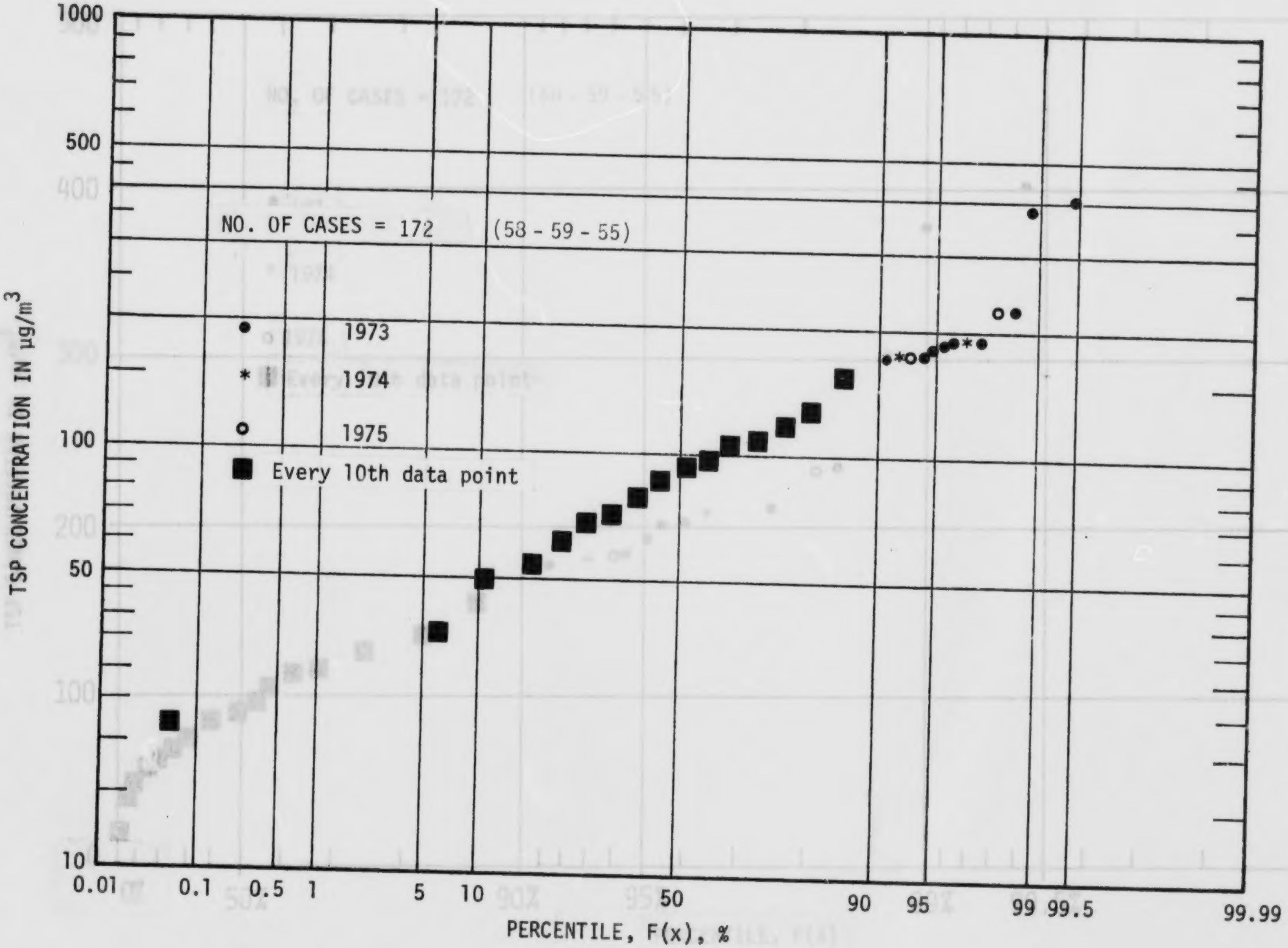


Figure A-1a. Log-normal distribution plot of TSP data at Forrest City, Arkansas (#040900001F01).

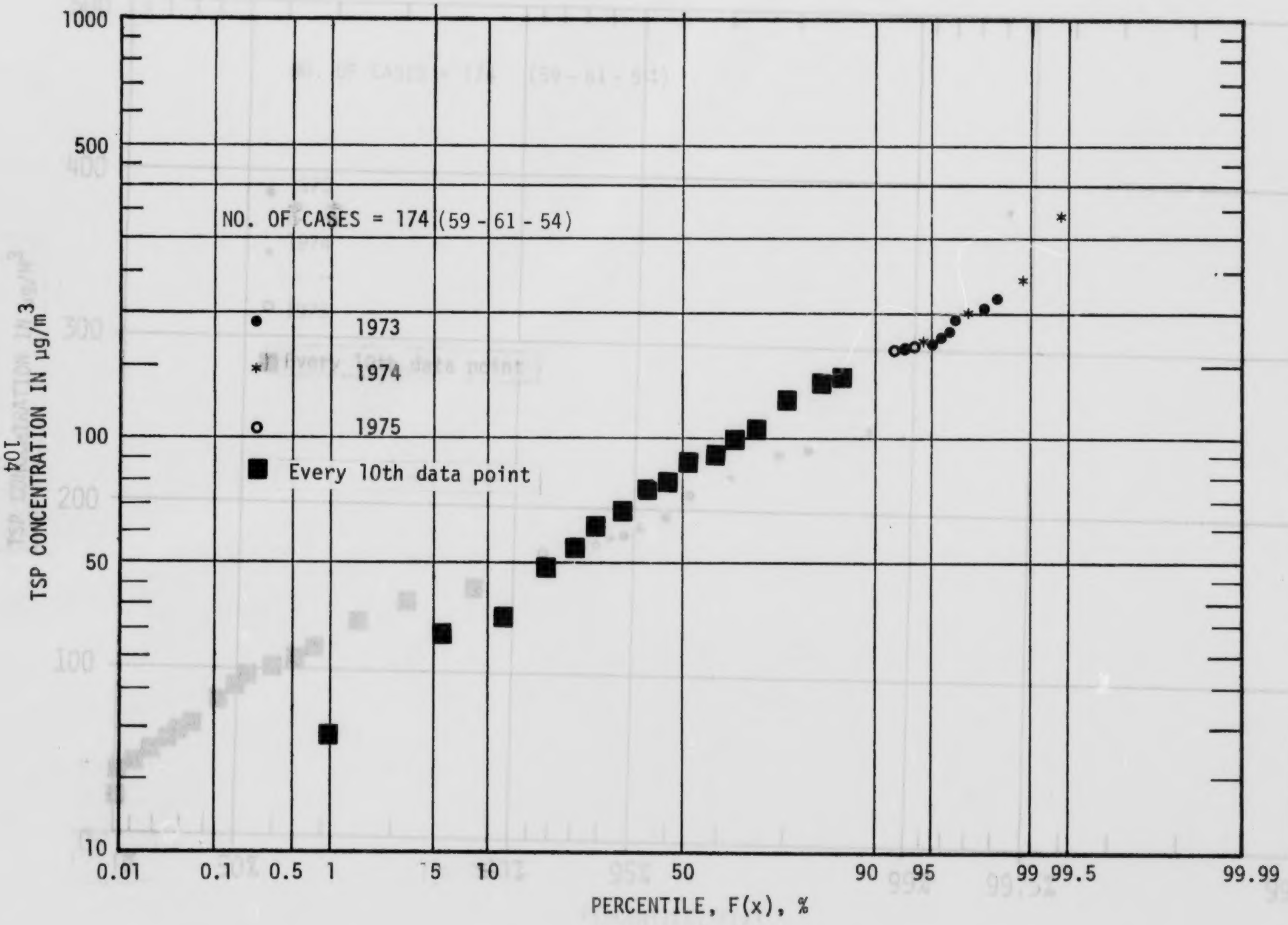


Figure A-2a. Log-normal distribution plot of TSP data at Helena, Arkansas (#041080001F01).

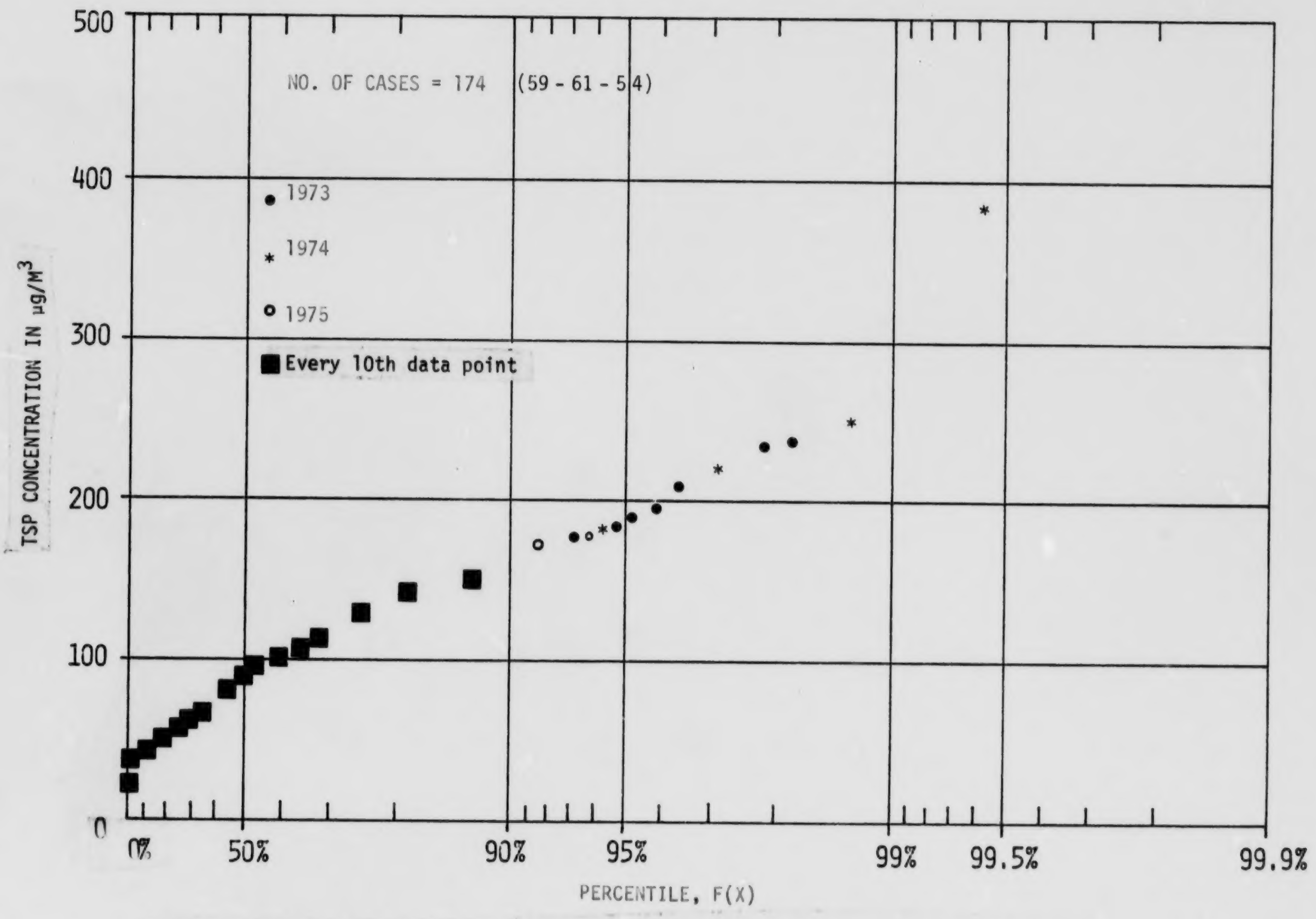


Figure A-2b. Exponential Distribution Plot of TSP Data at Helena, ARK (#041080001F01)

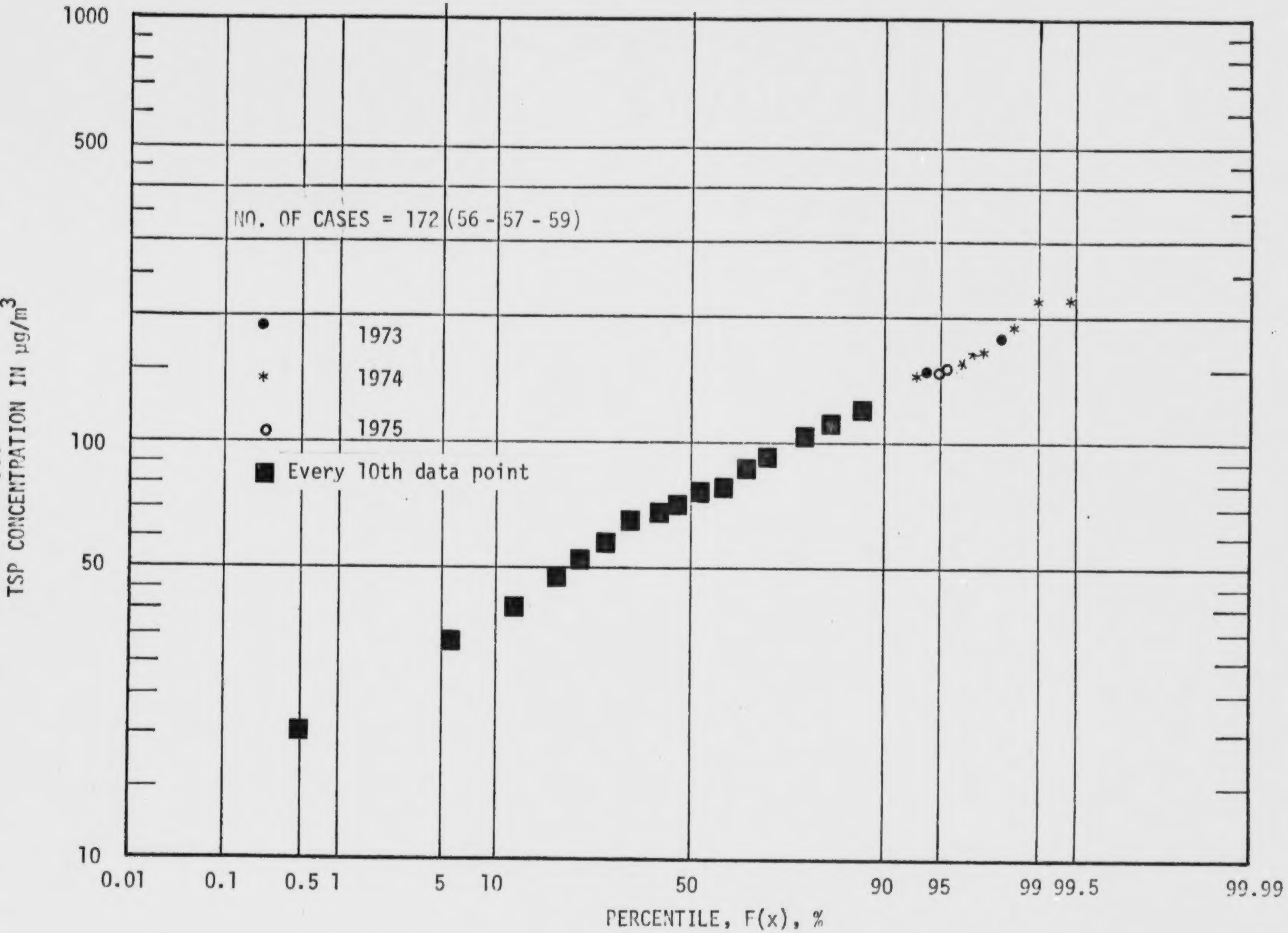


Figure A-3a. Log-normal distribution plot of TSP data at Jonesboro, Arkansas (#041320001F01).

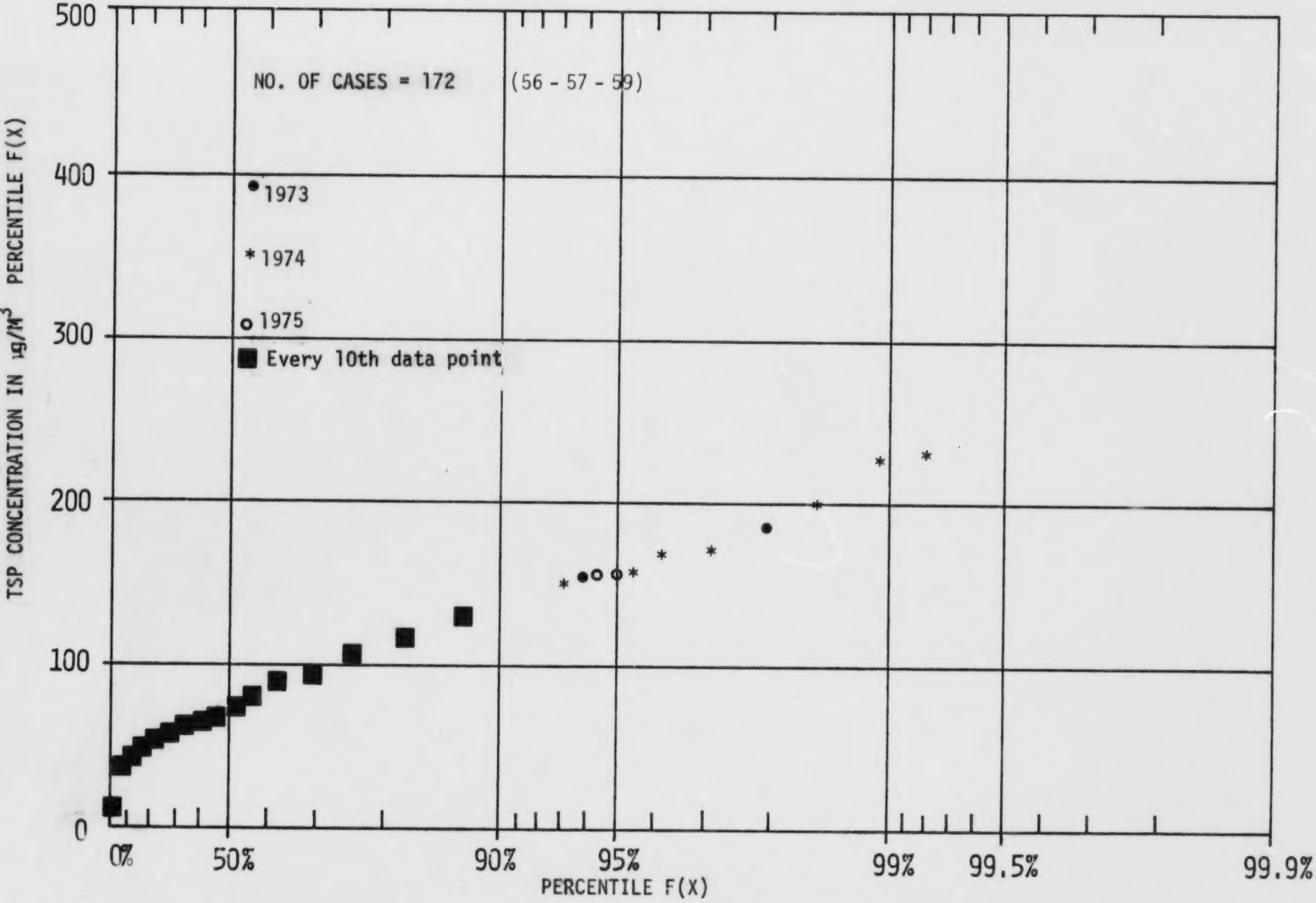


Figure A-3b. Exponential Distribution Plot of TSP Data at Jonesboro, Arkansas (#041320001F01)

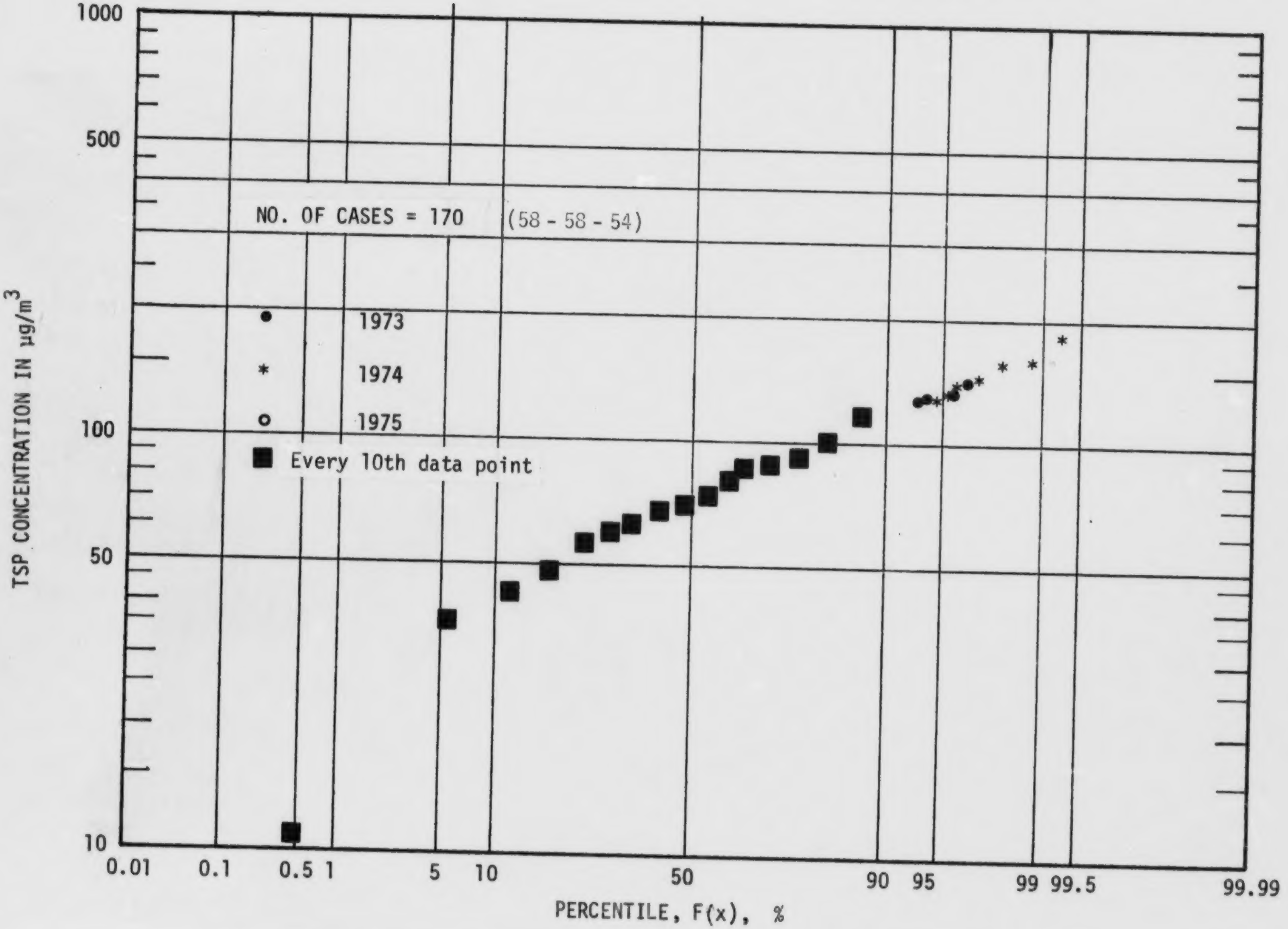


Figure A-4a. Log-normal distribution plot of TSP data at Springdale, Arkansas (#042480001F01).

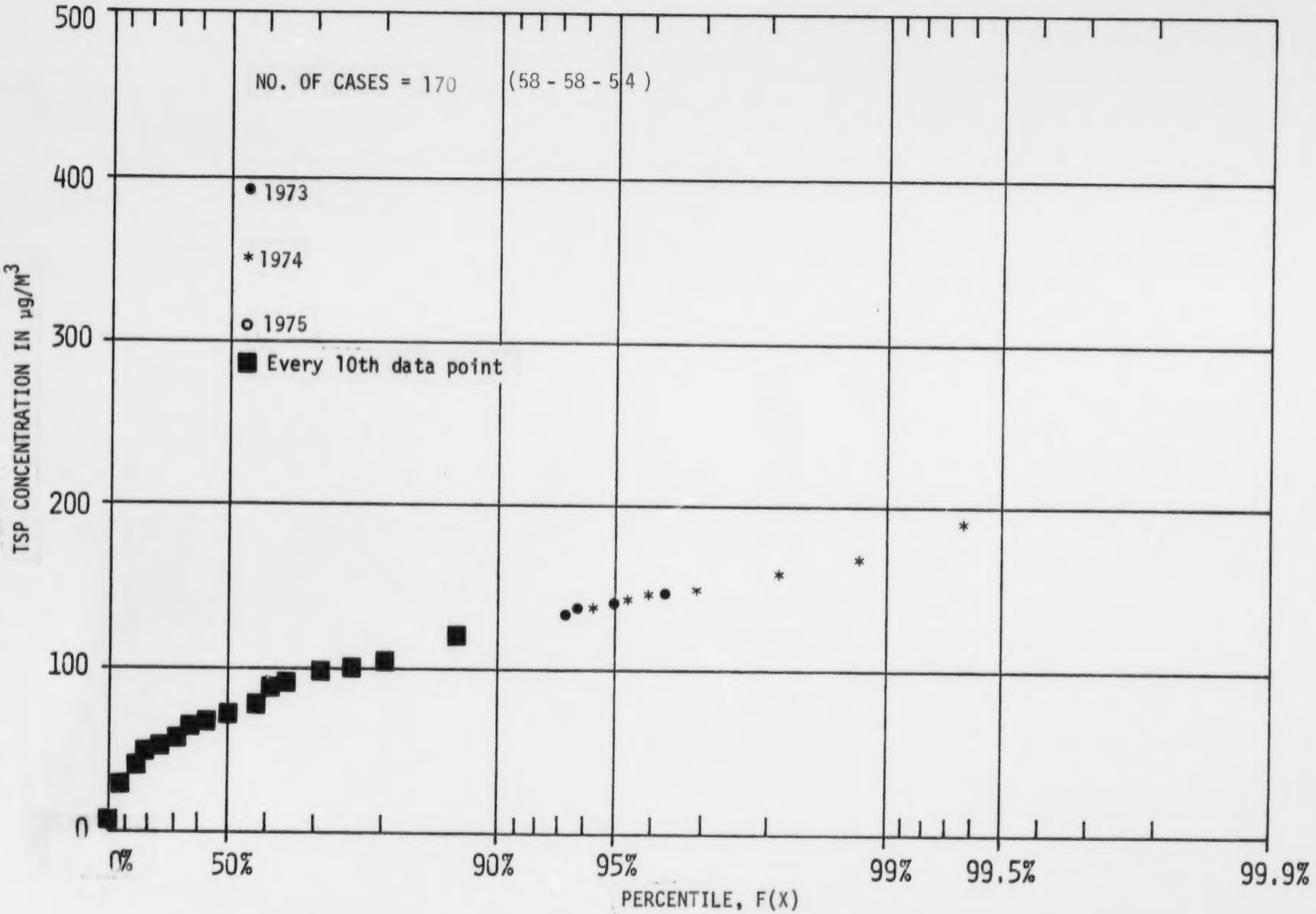


Figure A-4b. Exponential Distribution Plot of TSP Data at Springdale, ARK. (#042480001F01)

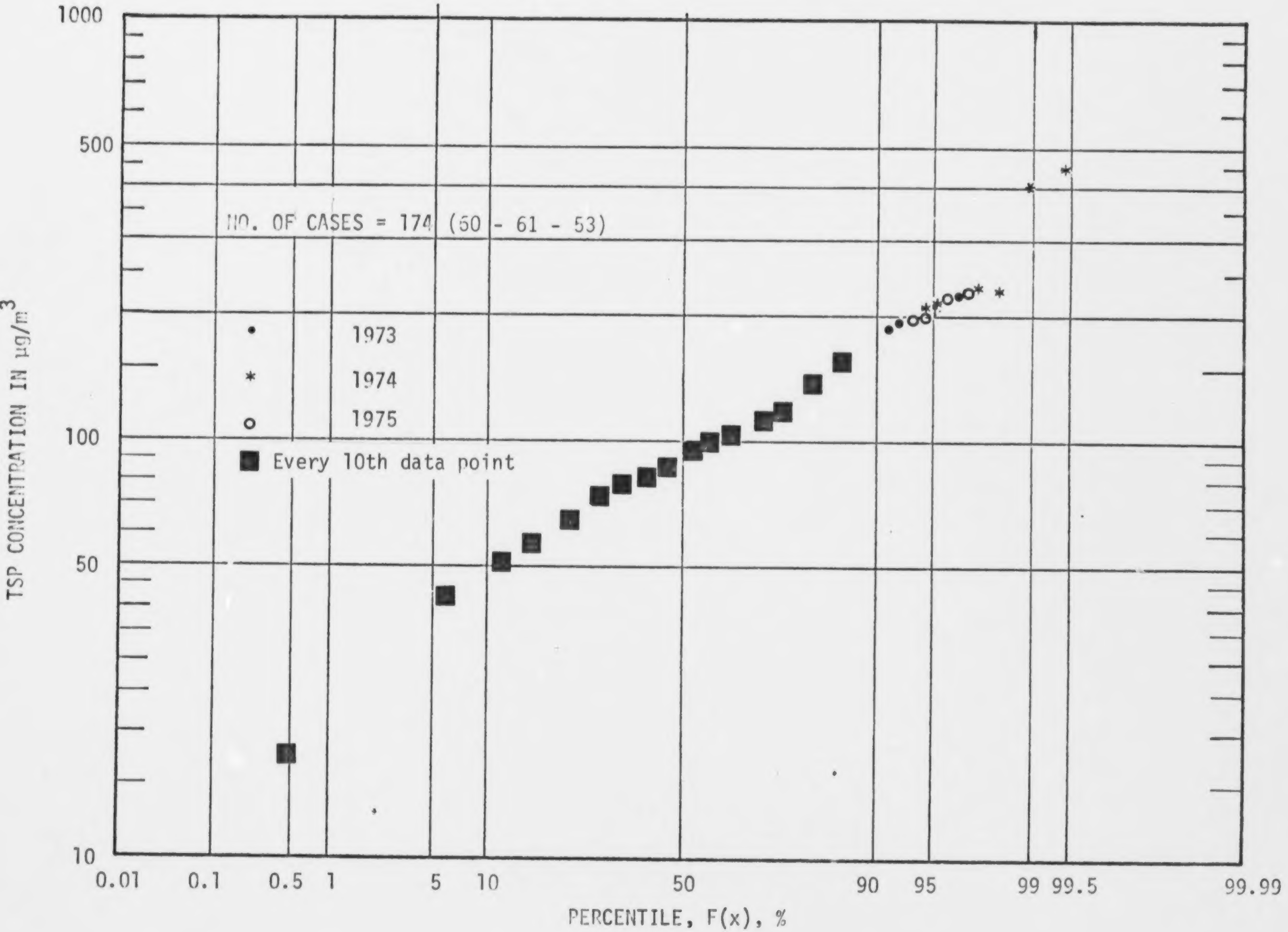


Figure A-5a Log-normal distribution plot of TSP data at Stuttgart, Arkansas (#042540001F01).

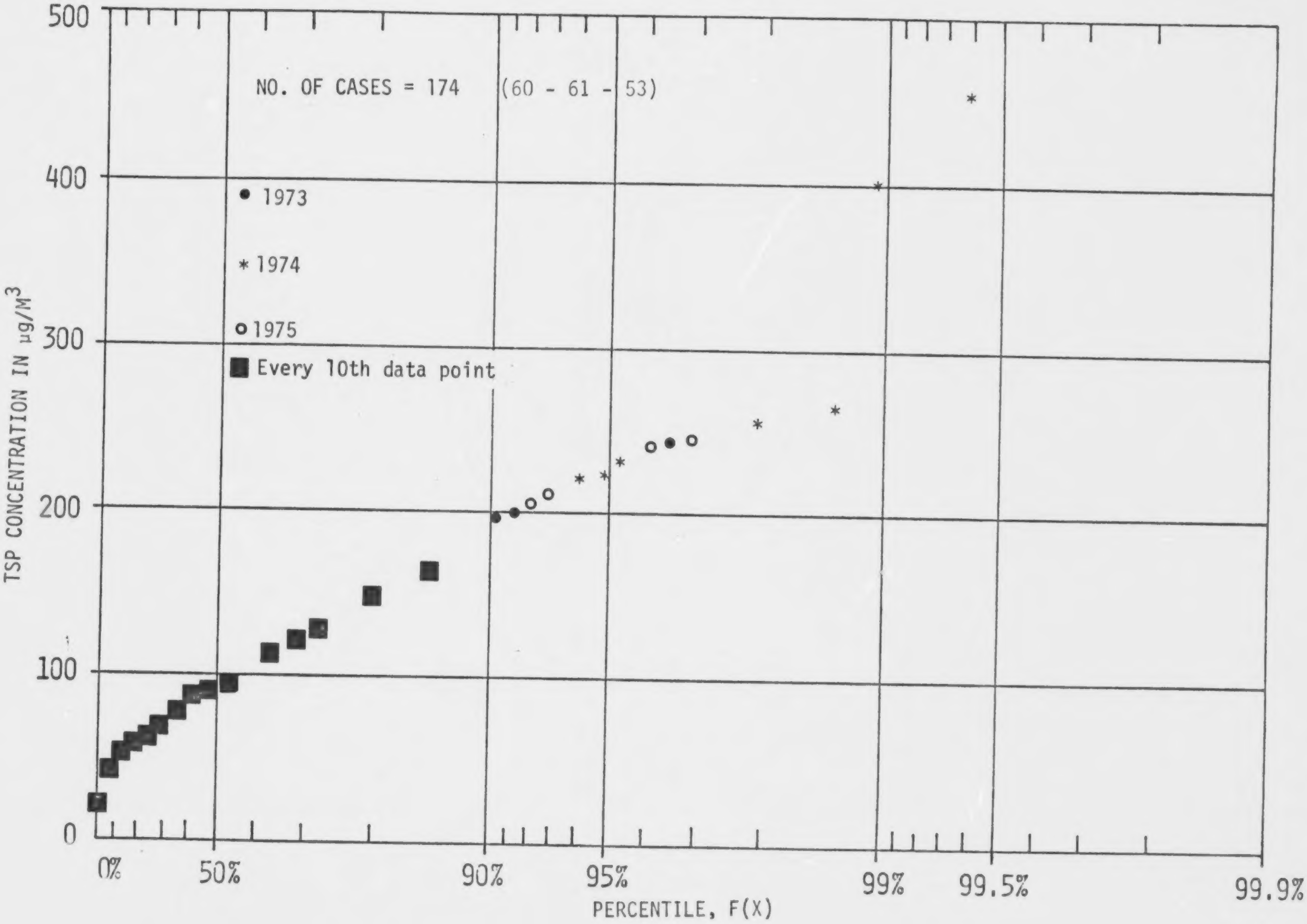


Figure A-5b. Exponential Distribution Plot of TSP Data at Stuttgart, ARK (#042540001F01)

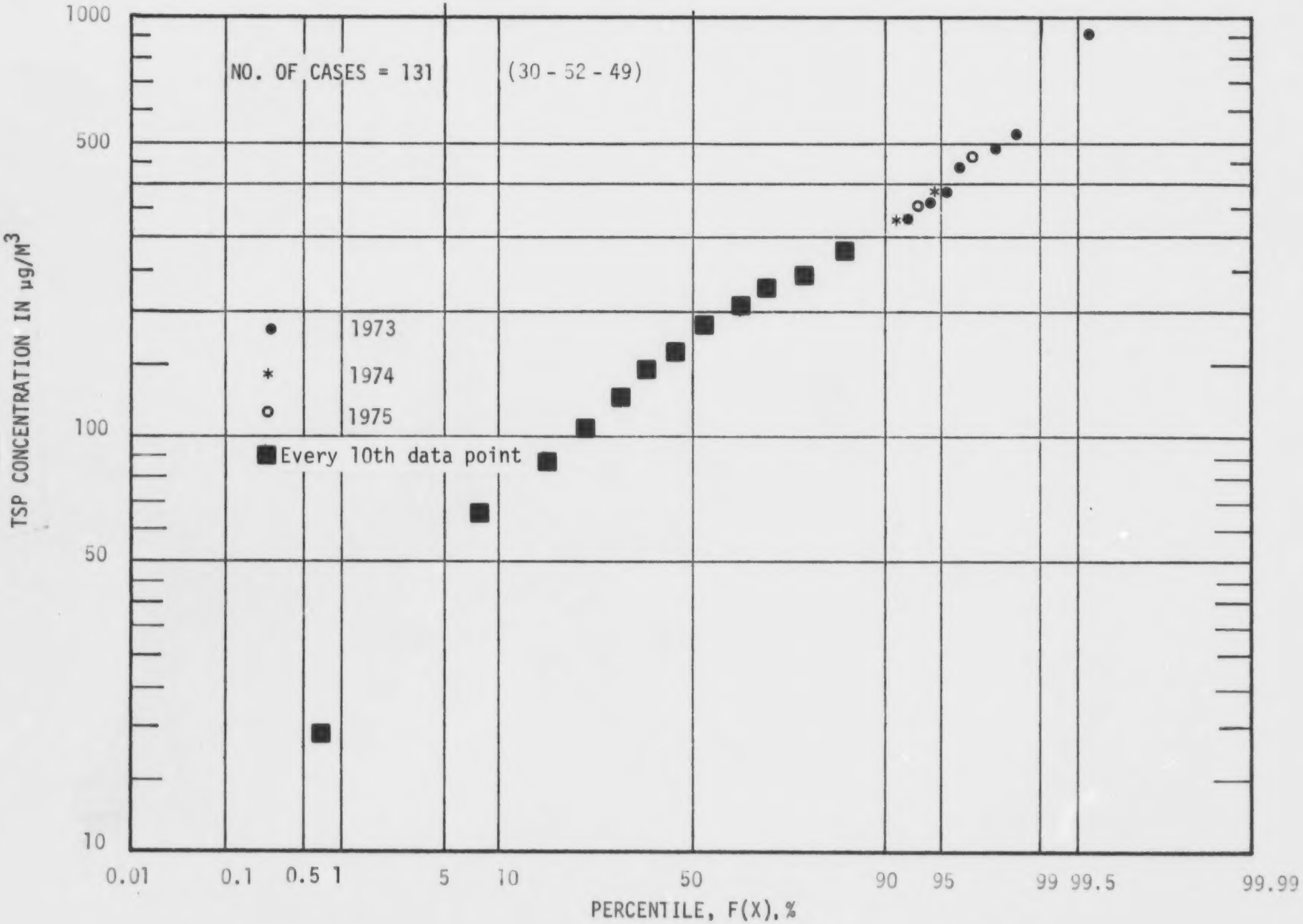


Figure A-6a. Log-Normal Distribution Plot of TSP Data at Stuttgart, ARK. (#042540005F05)

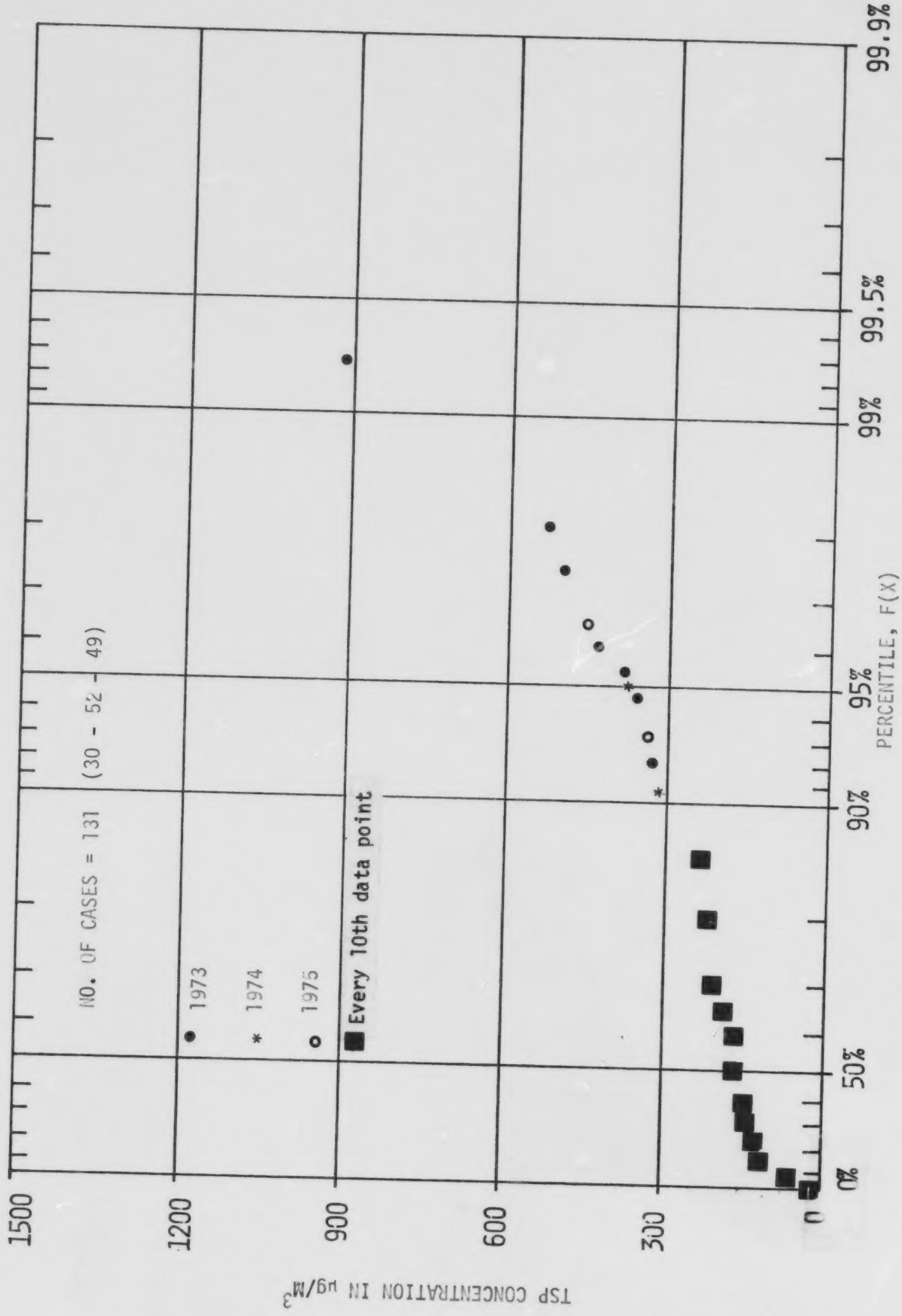


Figure A-6b. Exponential Distribution Plot of TSP Data at Stuttgart, ARK (#042540005F05)

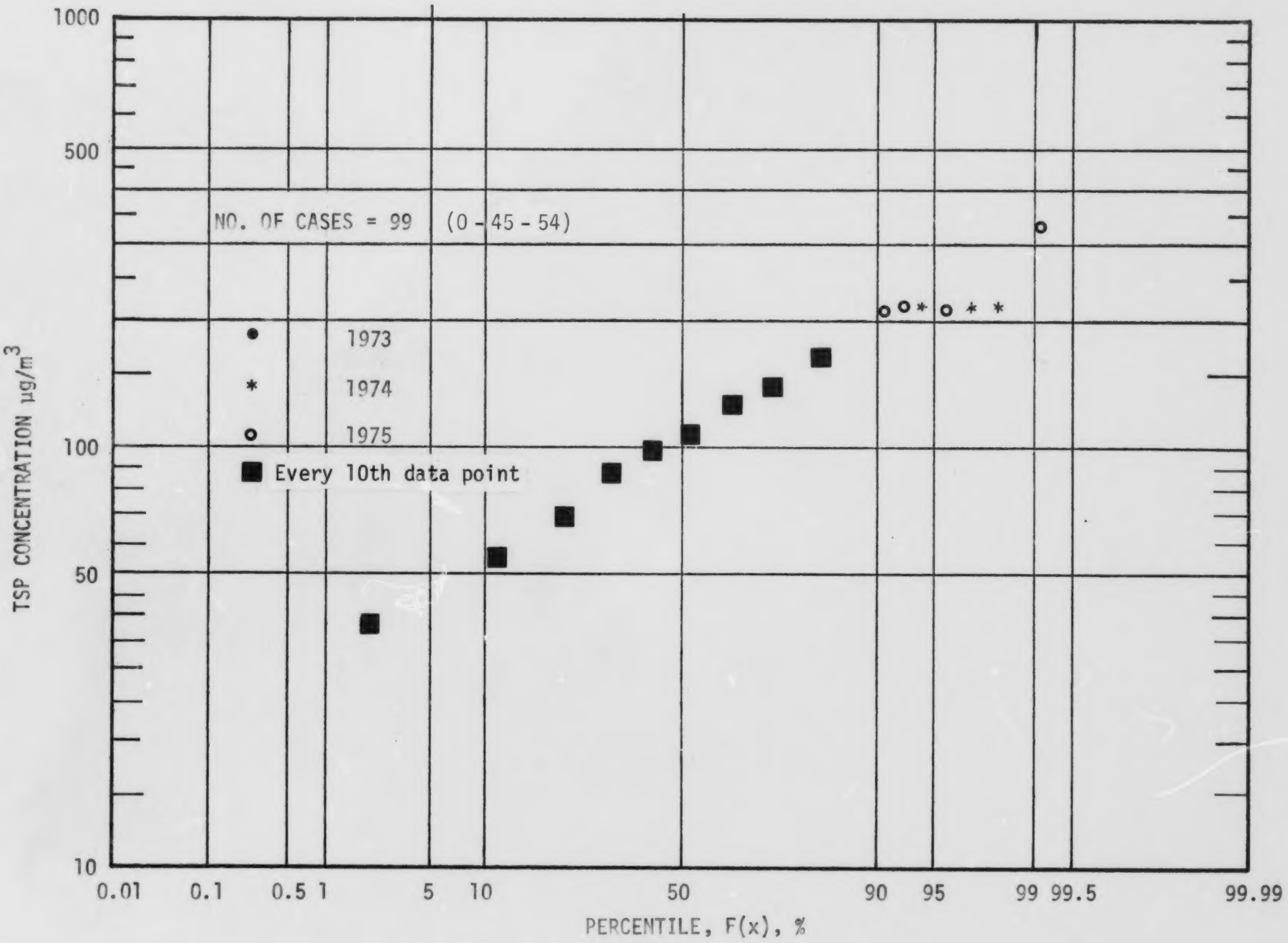


Figure A-7a. Log-normal distribution plot of TSP data at Albuquerque, New Mexico (#320040003H01).

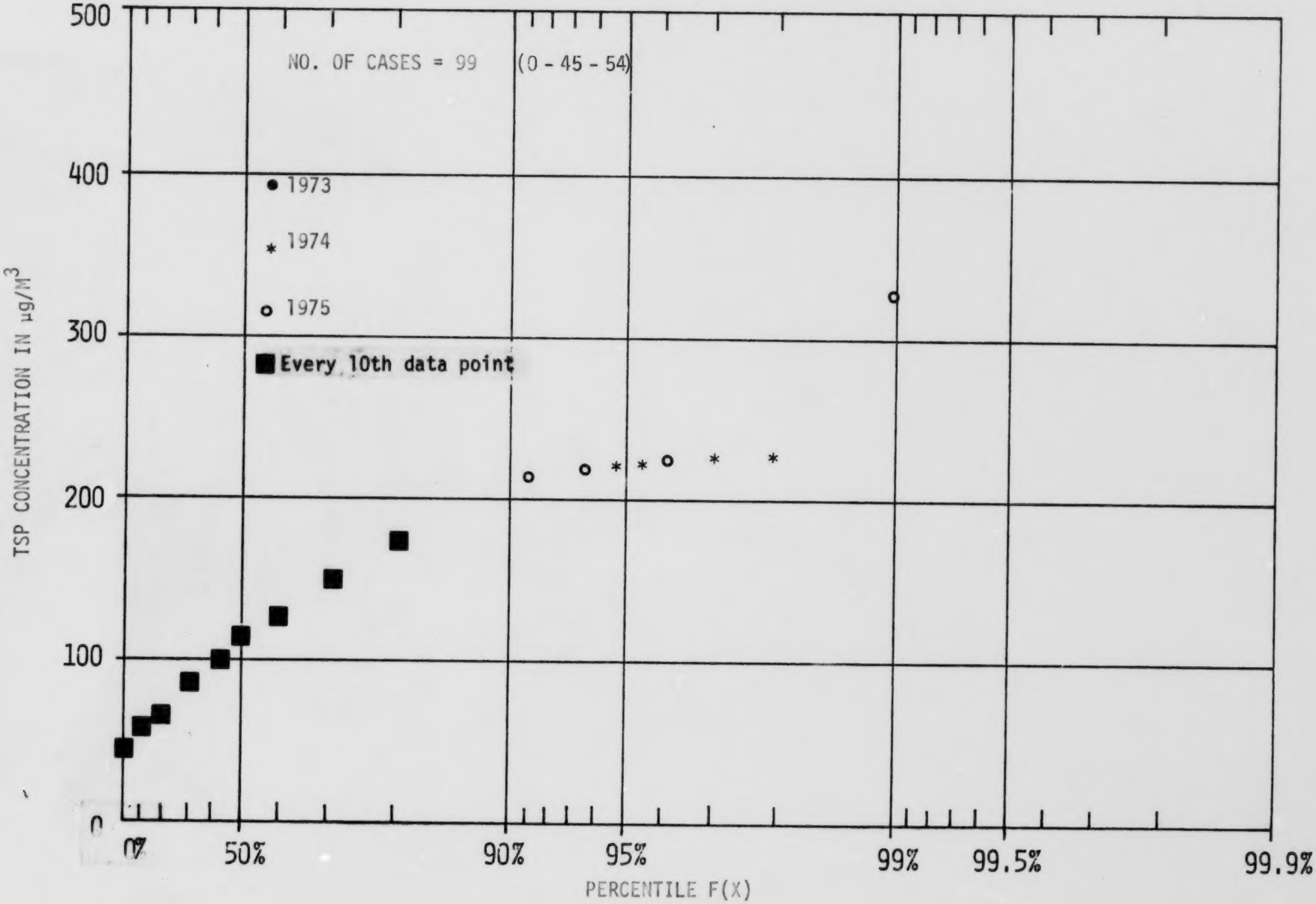


Figure A-7b. Exponential Distribution Plot of TSP Data at Albuquerque, NM (#320040003H01)

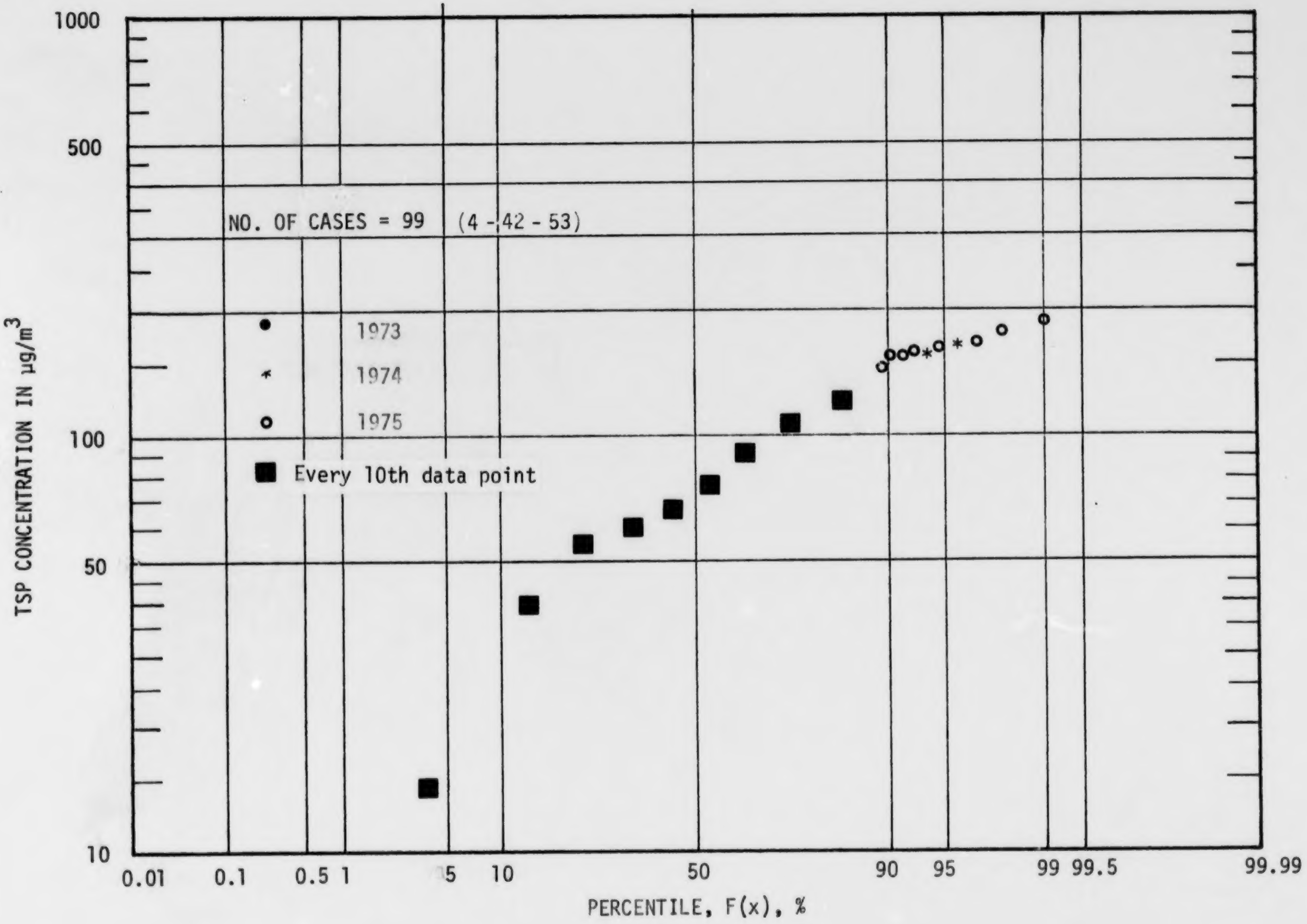


Figure A-8a. Log-normal distribution plot of TSP data at Albuquerque, New Mexico (#320040008H02).

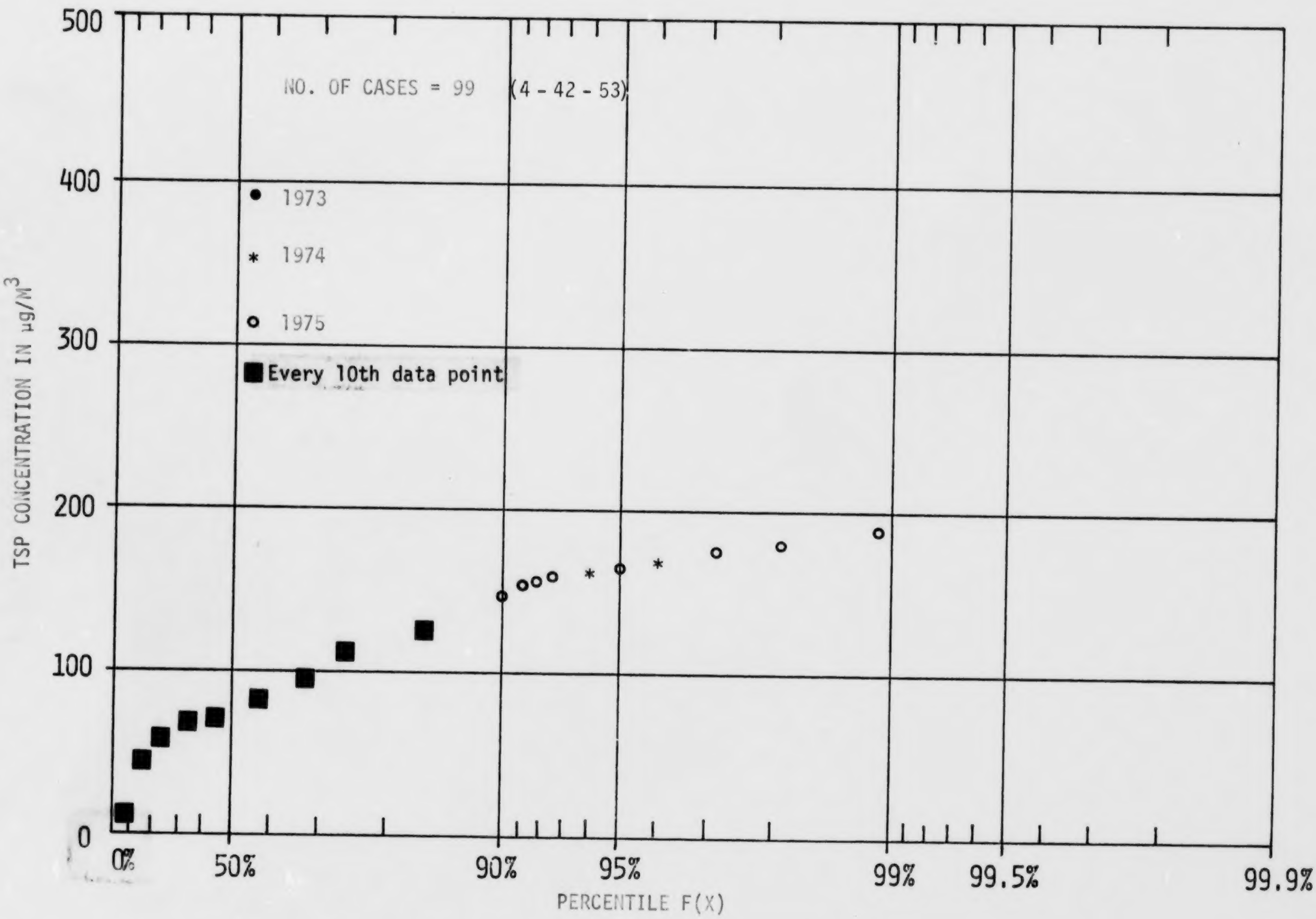


Figure A-8b. Exponential Distribution Plot of TSP Data at Albuquerque, NM (#320040008H02)

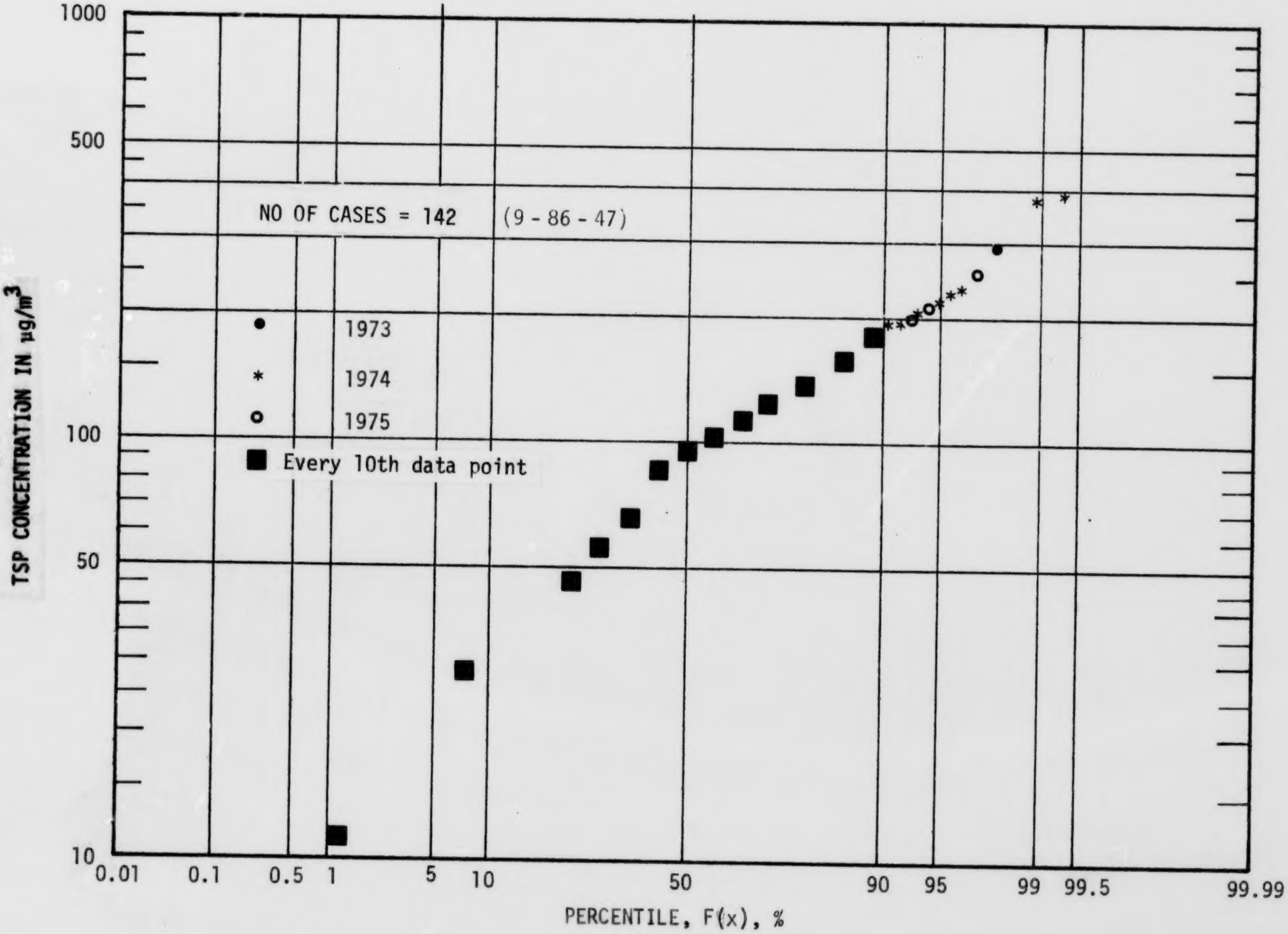


Figure A-9a. Log-normal distribution plot of TSP data at Dona Ana, N.M. (#320340001F02).

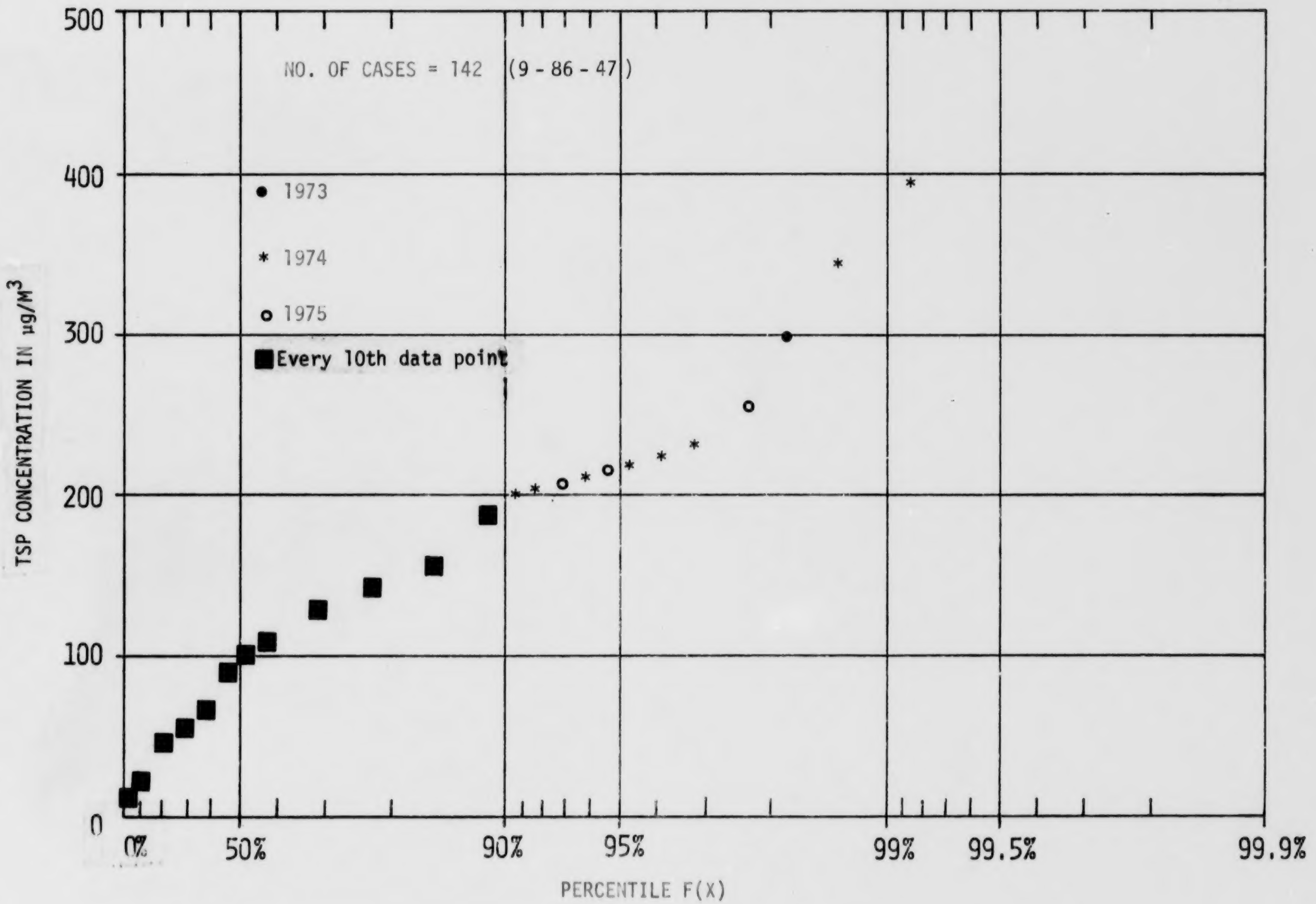


Figure A-9b. Exponential Distribution Plot of TSP Data at Dona Ana CO. (#320340001F02)

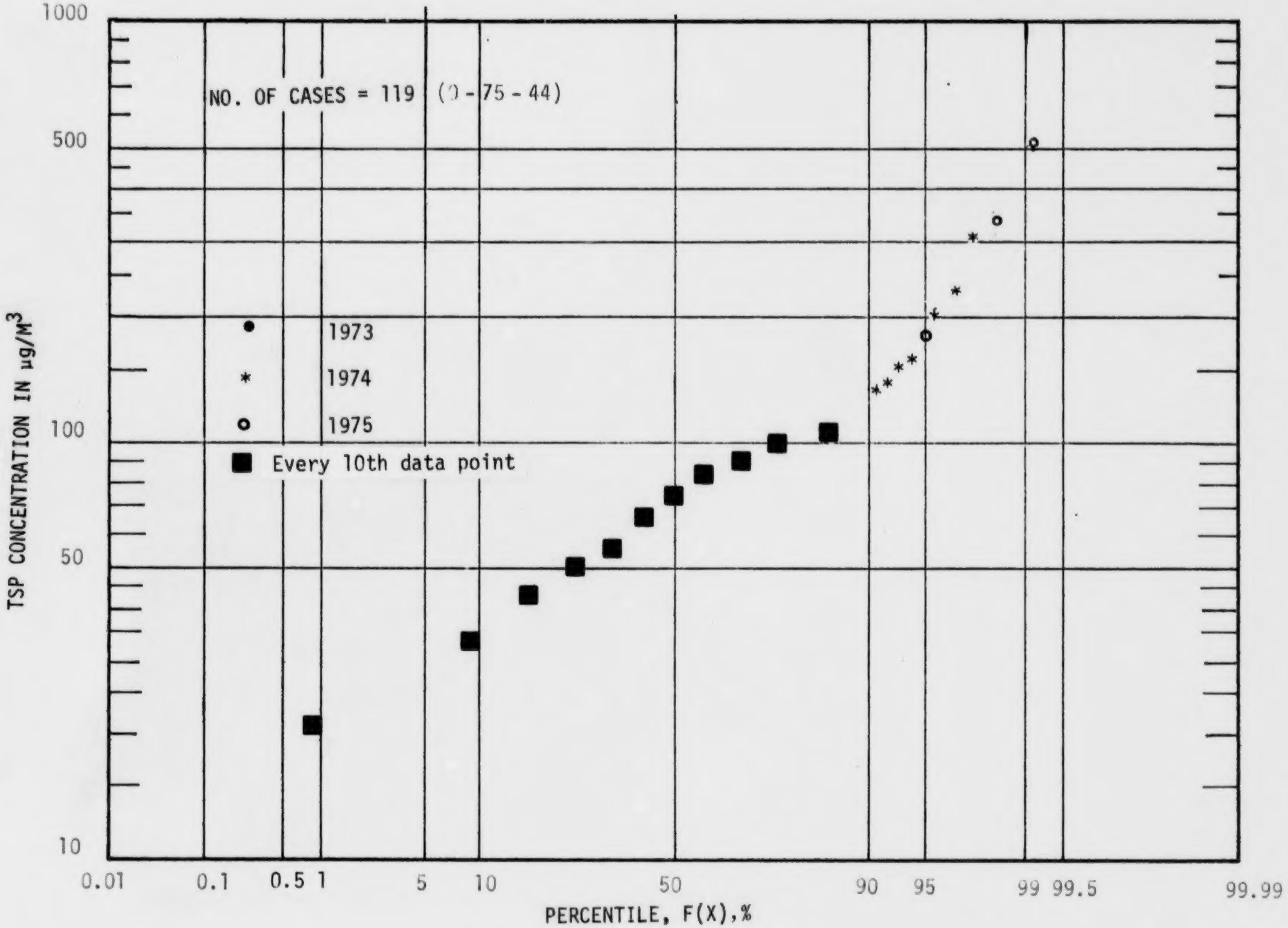


Figure A-10a. Log-Normal Distribution Plot of TSP Data at Las Cruces, NM (#320580001F01)

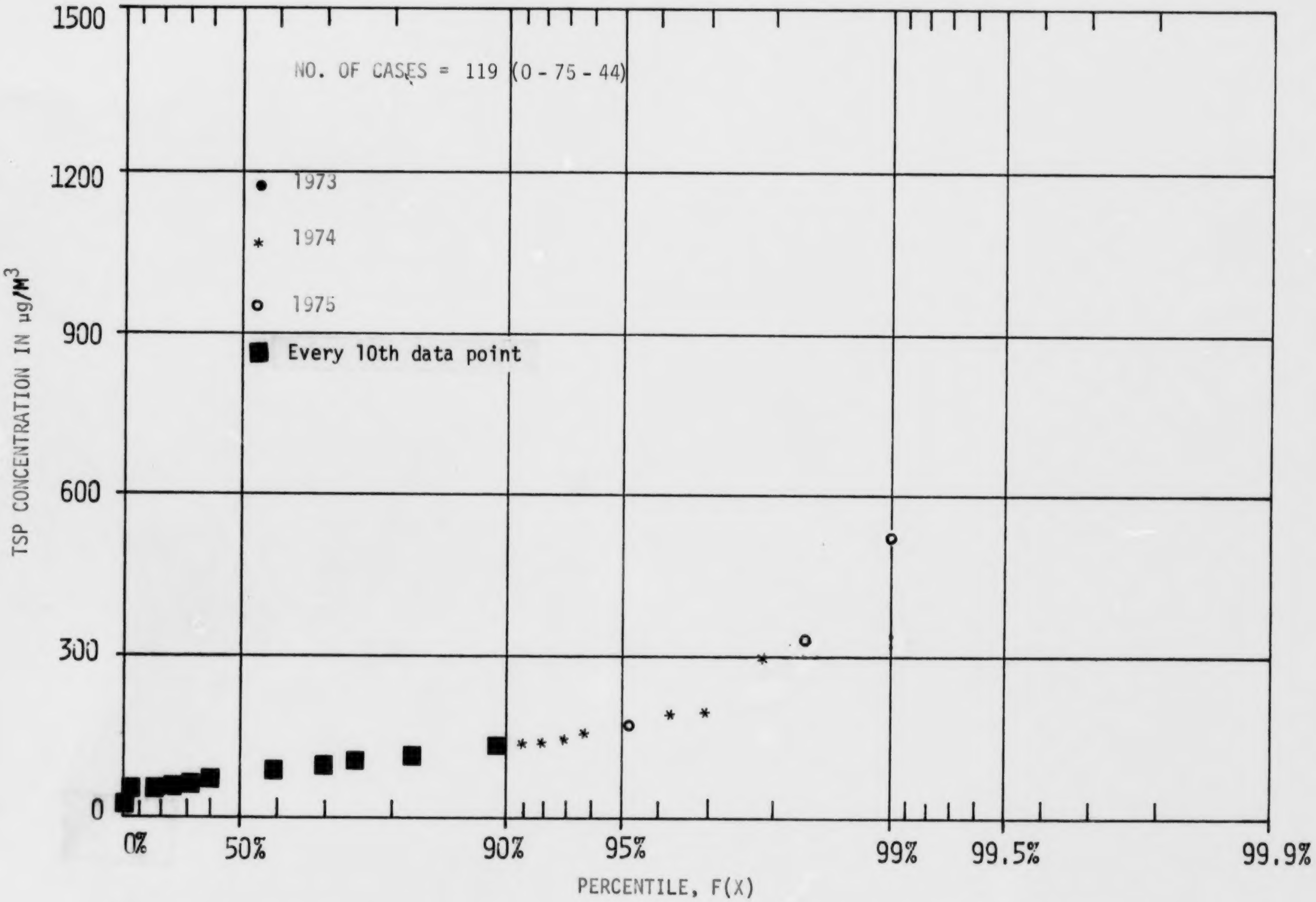


Figure A-10b. Exponential Distribution Plot of TSP Data at Las Cruces, NM (#320580001F01)

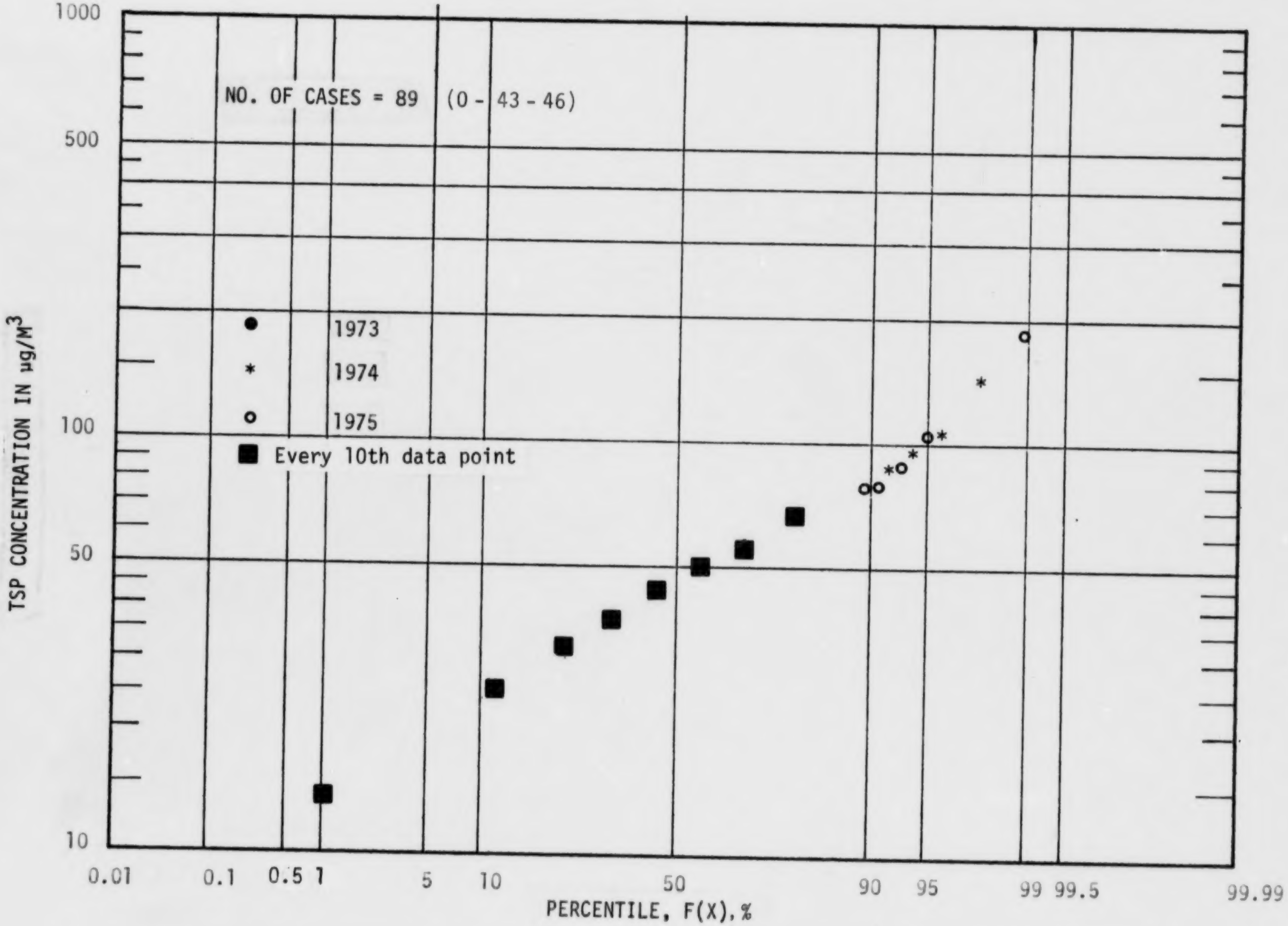


Figure A-11a. Log-Normal Distribution Plot of TSP Data at Raton, N.M. (#320900001F01)

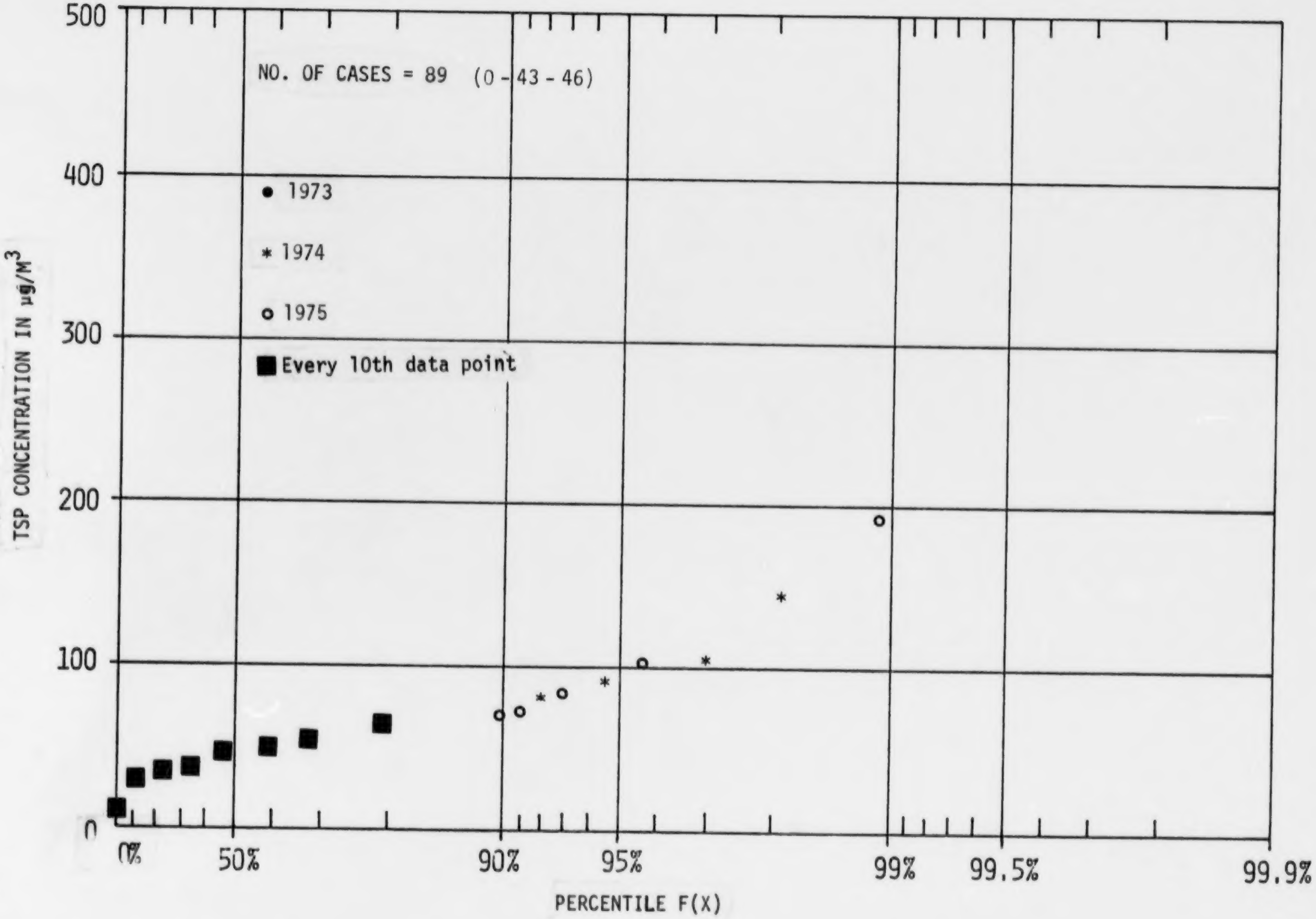


Figure A-11b. Exponential Distribution Plot of TSP Data at Raton, NM (#320900001F01)

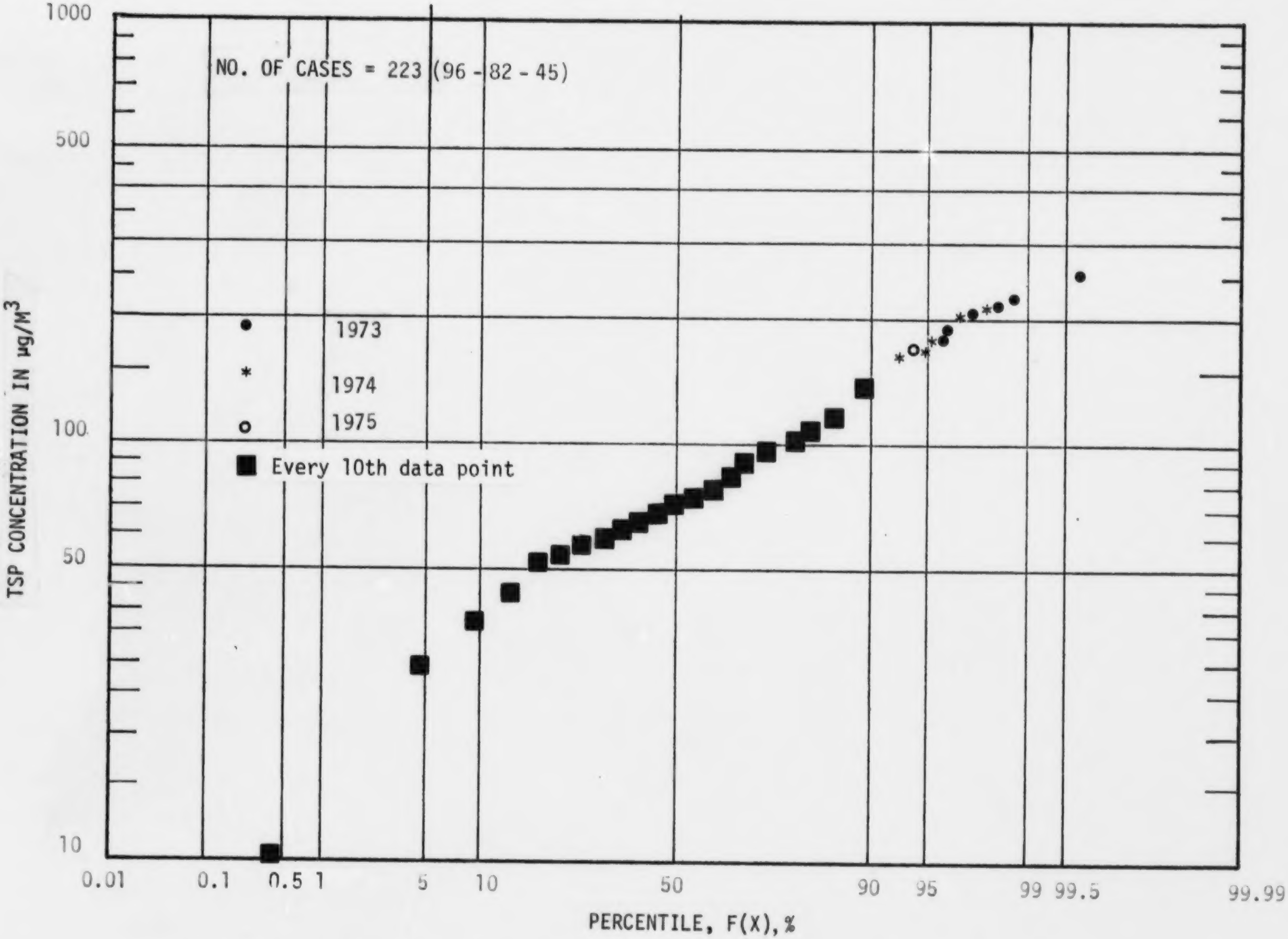


Figure A-12a. Log-Normal Distribution Plot of TSP Data at Bethany, OK (#370260014F01)

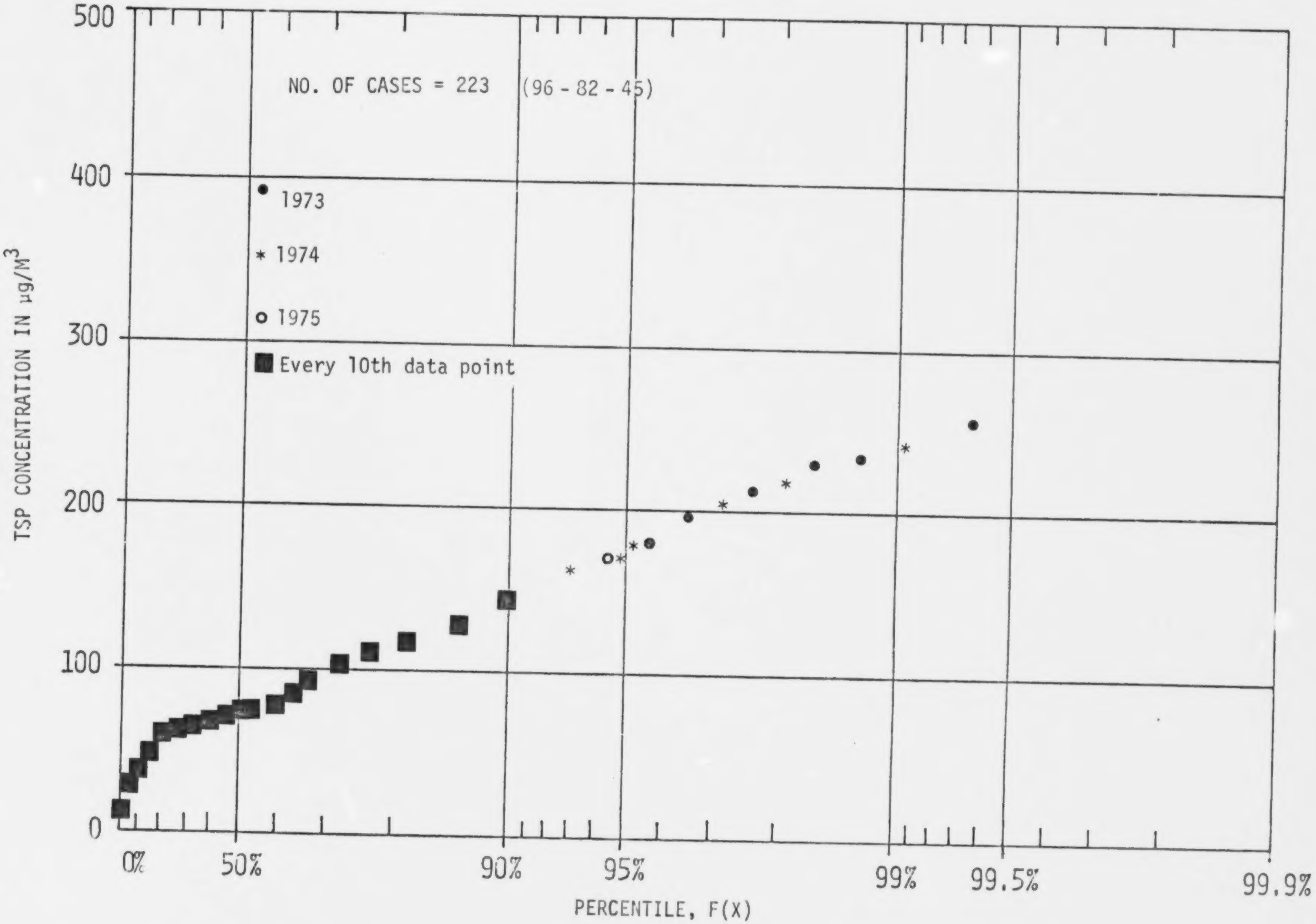


Figure A-12b. Exponential Distribution Plot of TSP Data at Bethany, OK (#370260014F01)

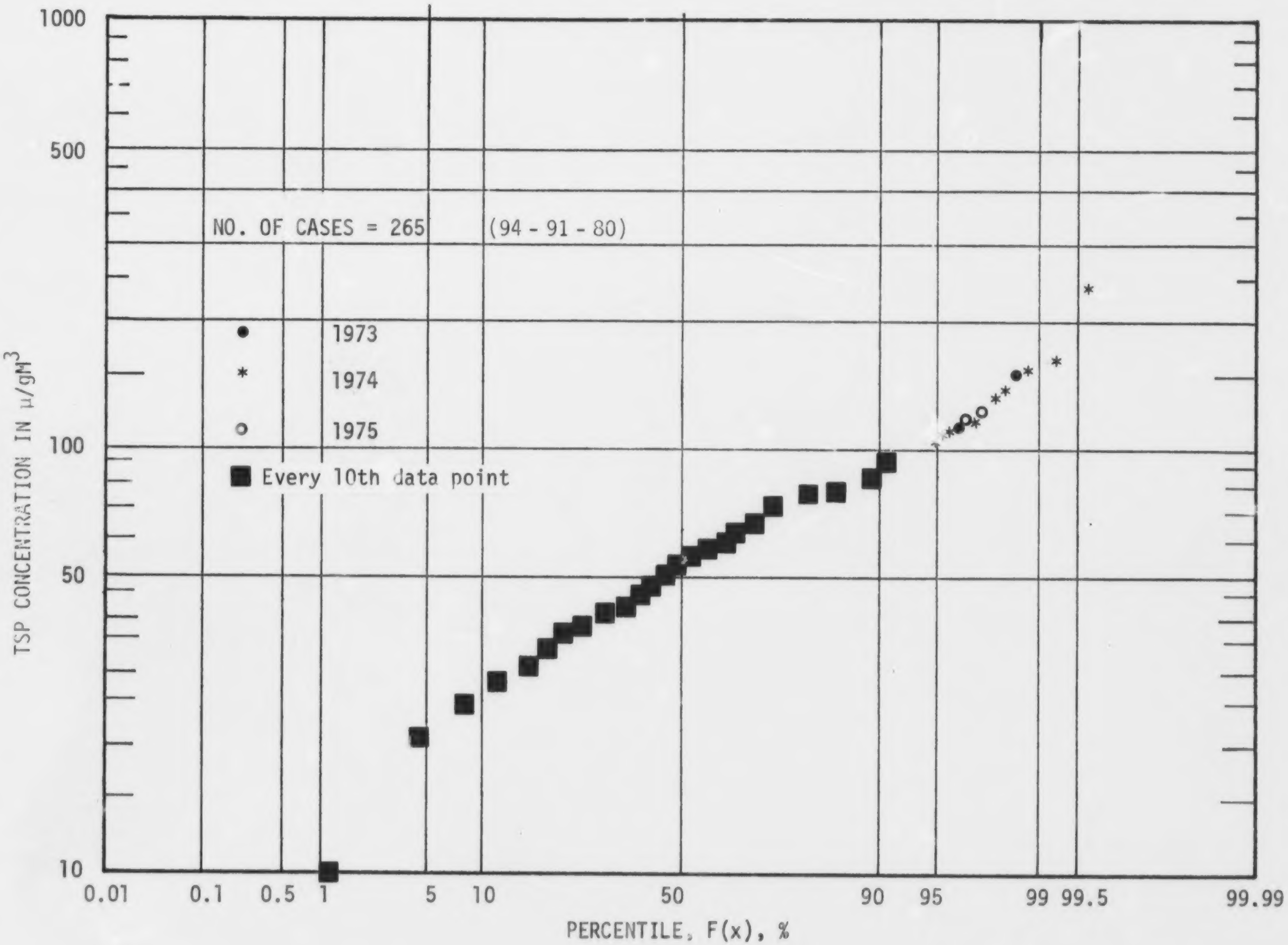


Figure A-13a. Log-normal distribution plot of TSP data at Muskogee, Oklahoma (#371980162F01).

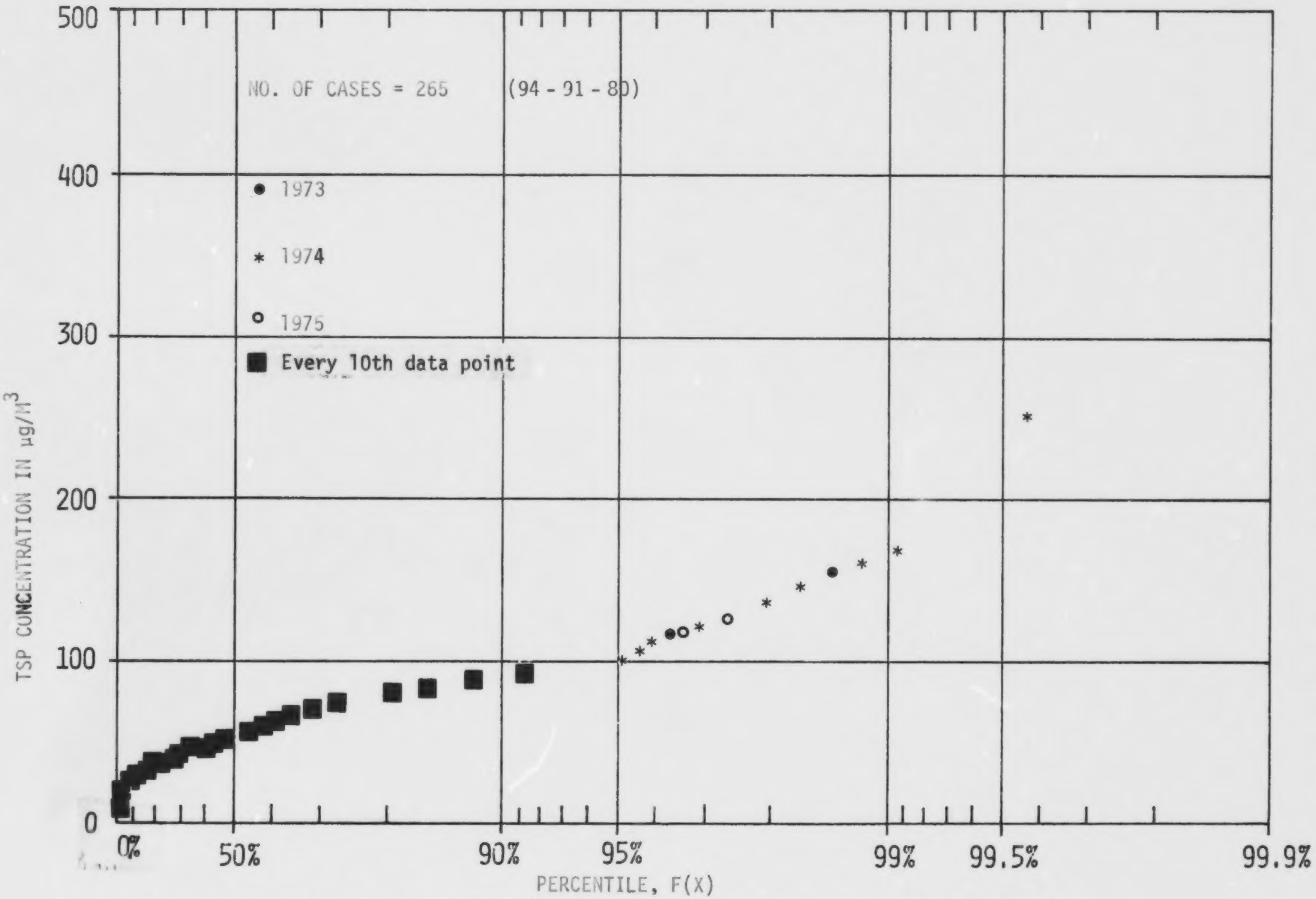


Figure A-13b. Exponential Distribution Plot of TSP Data at Muskogee, OK (#371980162F01)

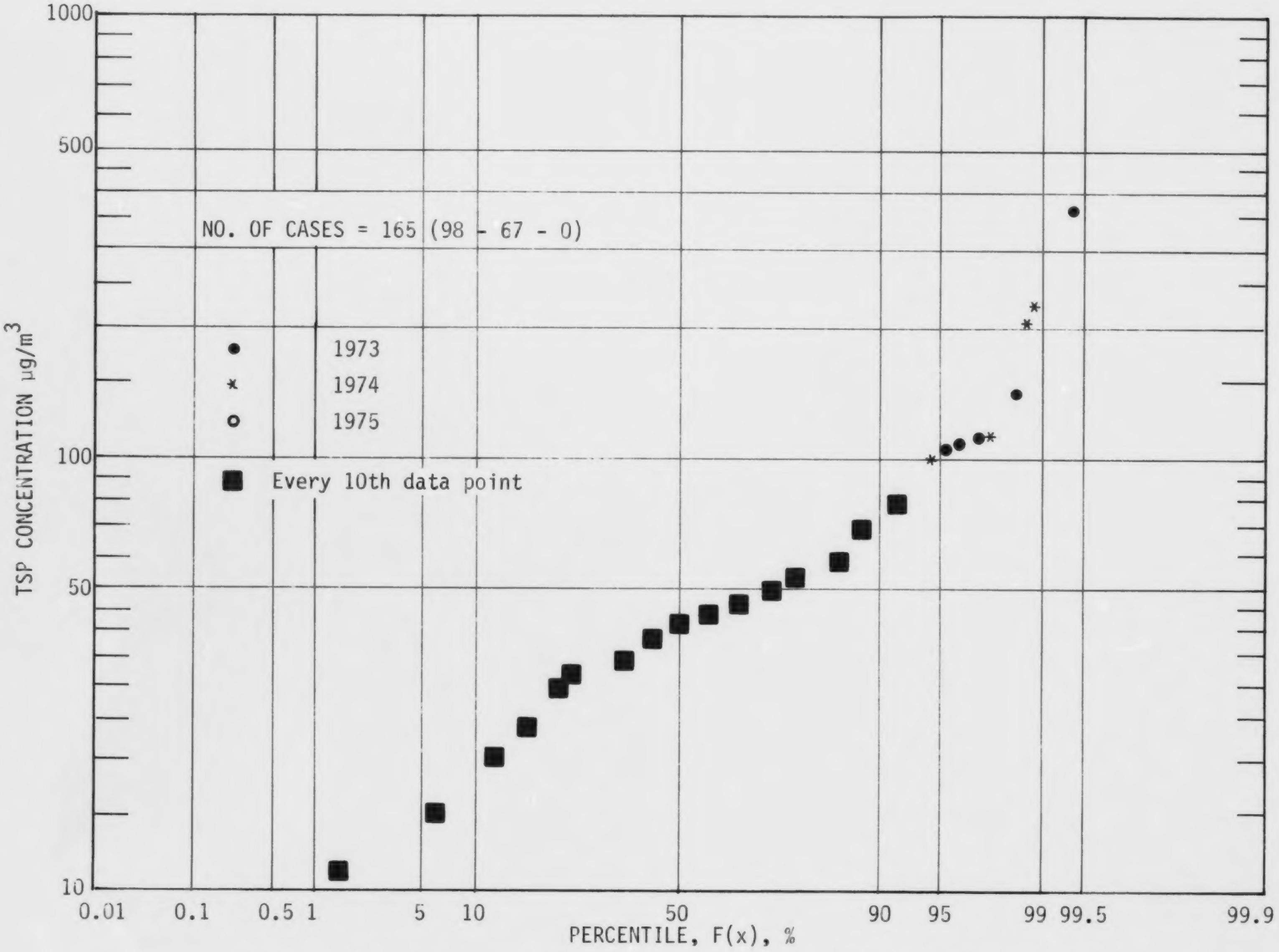


Figure A-14a. Log-normal distribution plot of TSP data at Oklahoma City, Oklahoma (#372200002F01).

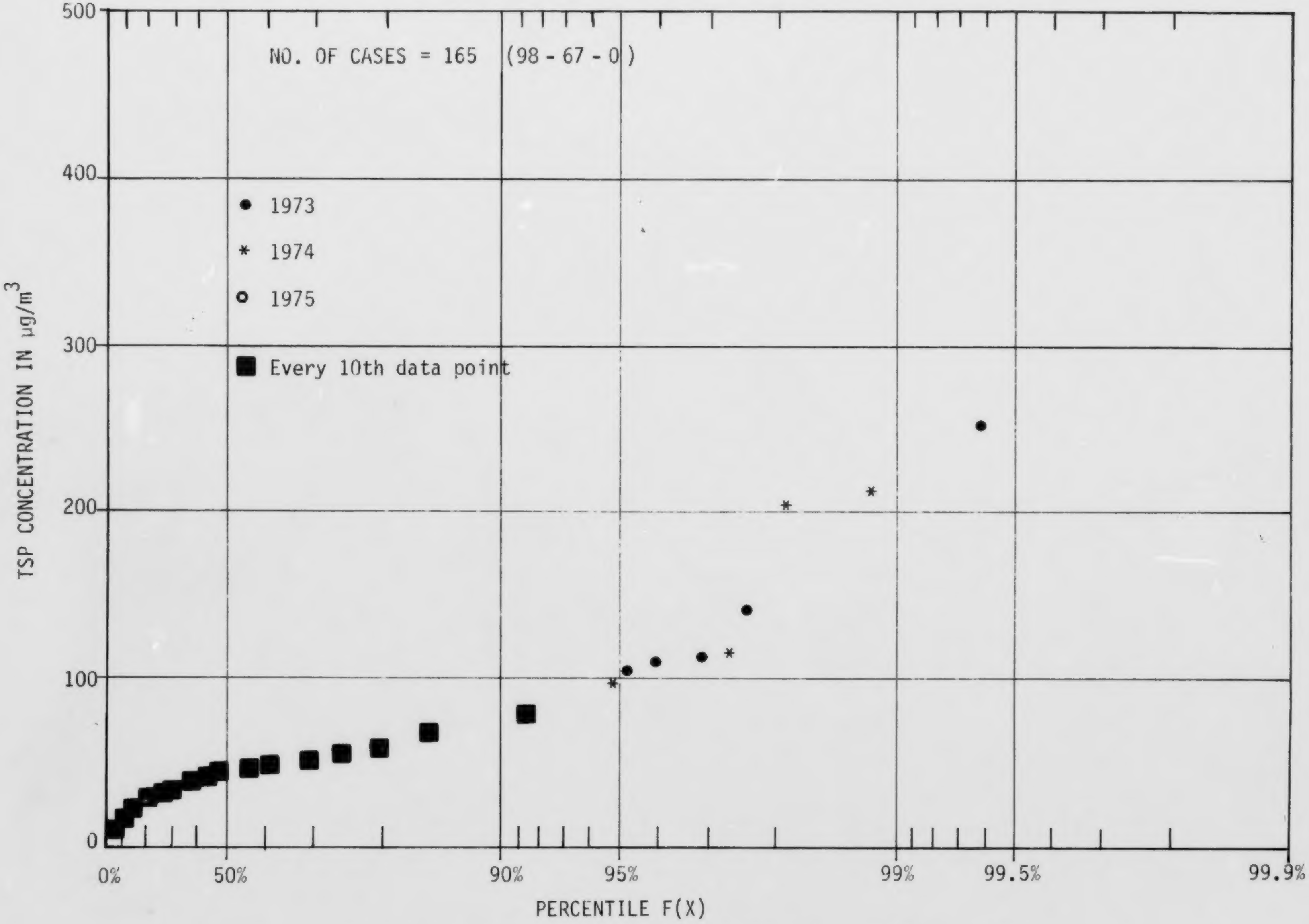


Figure A-14b. Exponential Distribution Plot of TSP Data at Oklahoma City, Oklahoma (#372200002F01).

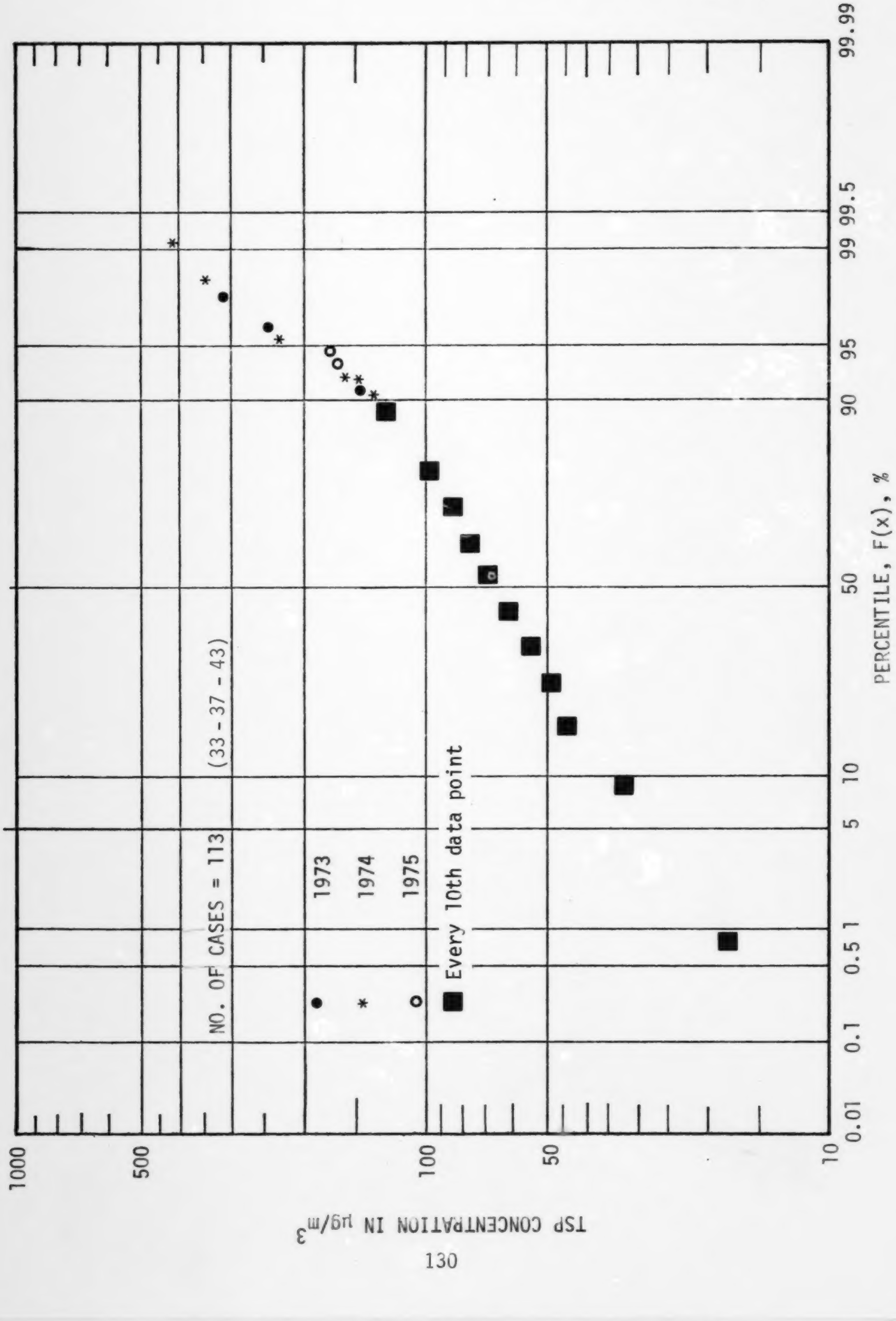


Figure A-15a. Log-normal distribution plot of TSP data at Roger Mills Co., Oklamoma (#372600900F01).

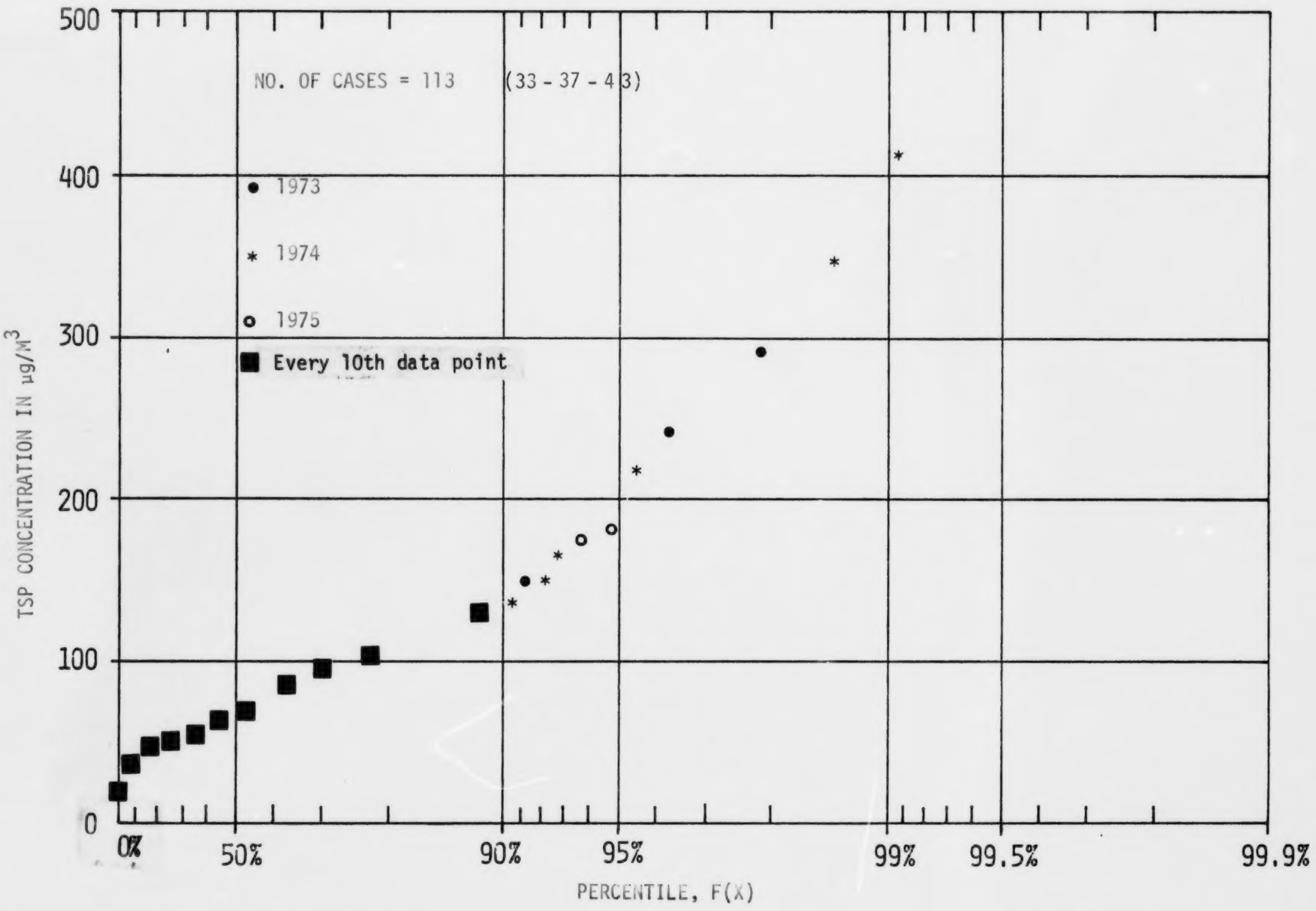


Figure A-15b. Exponential Distribution Plot of TSP Data at Roger Mills Co., OK (#372600900F01)

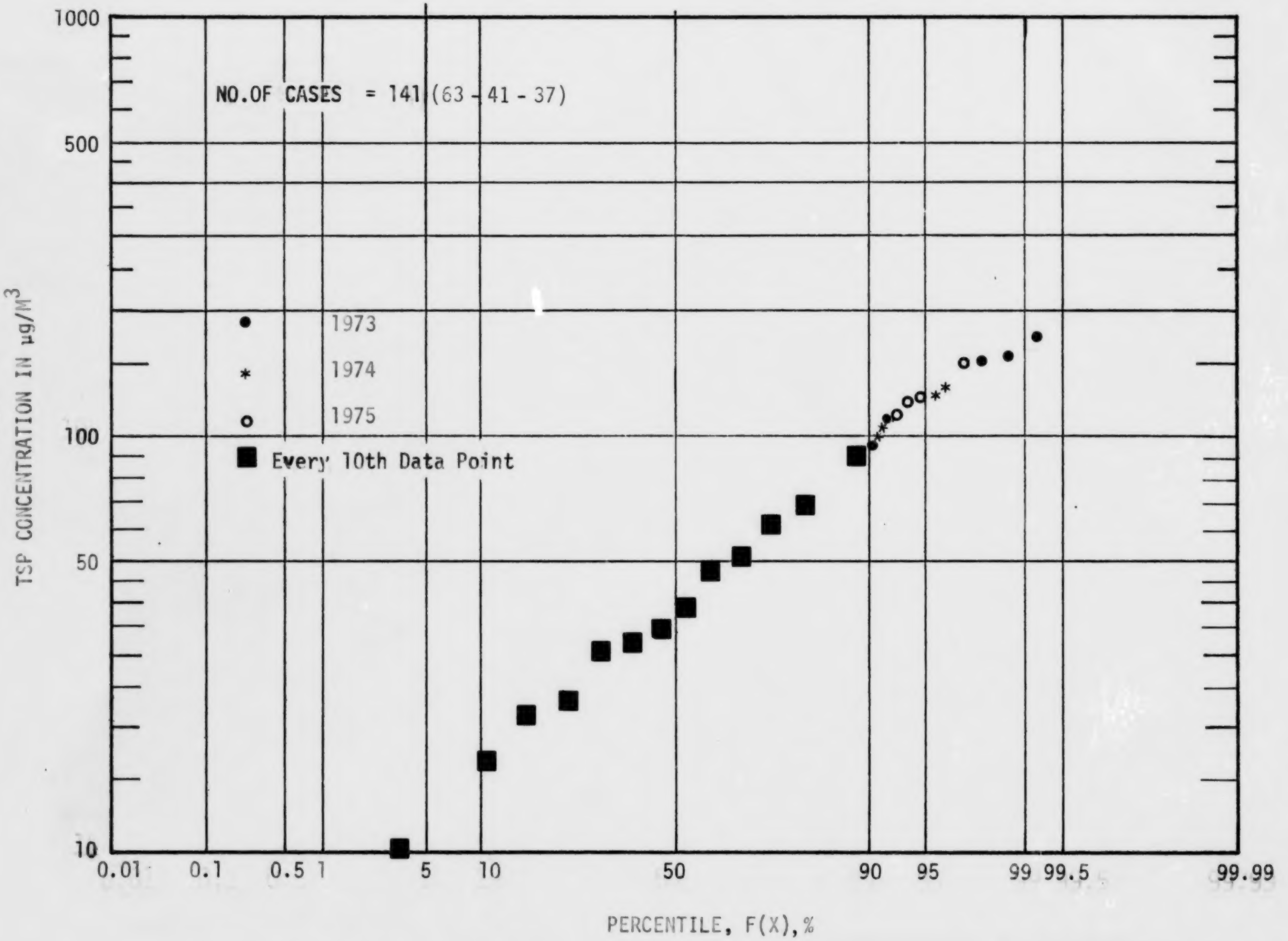


Figure A-16a. Log-Normal Distribution Plot of TSP Data at Sequoyah Co., OK (#372760471F01)

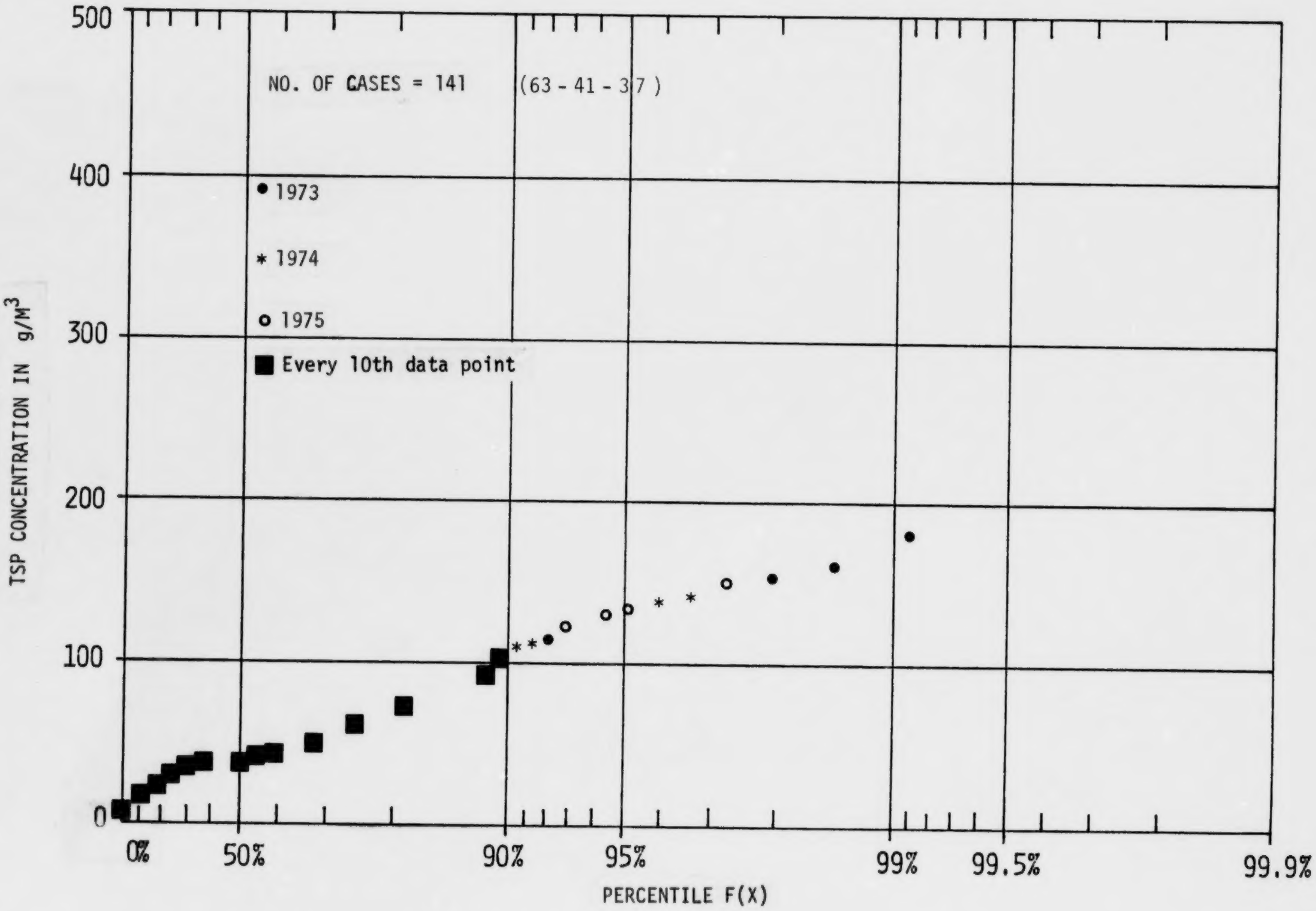


Figure A-16b. Exponential Distribution Plot of TSP Data at Sequoyah Co., OK (#372760471F01)

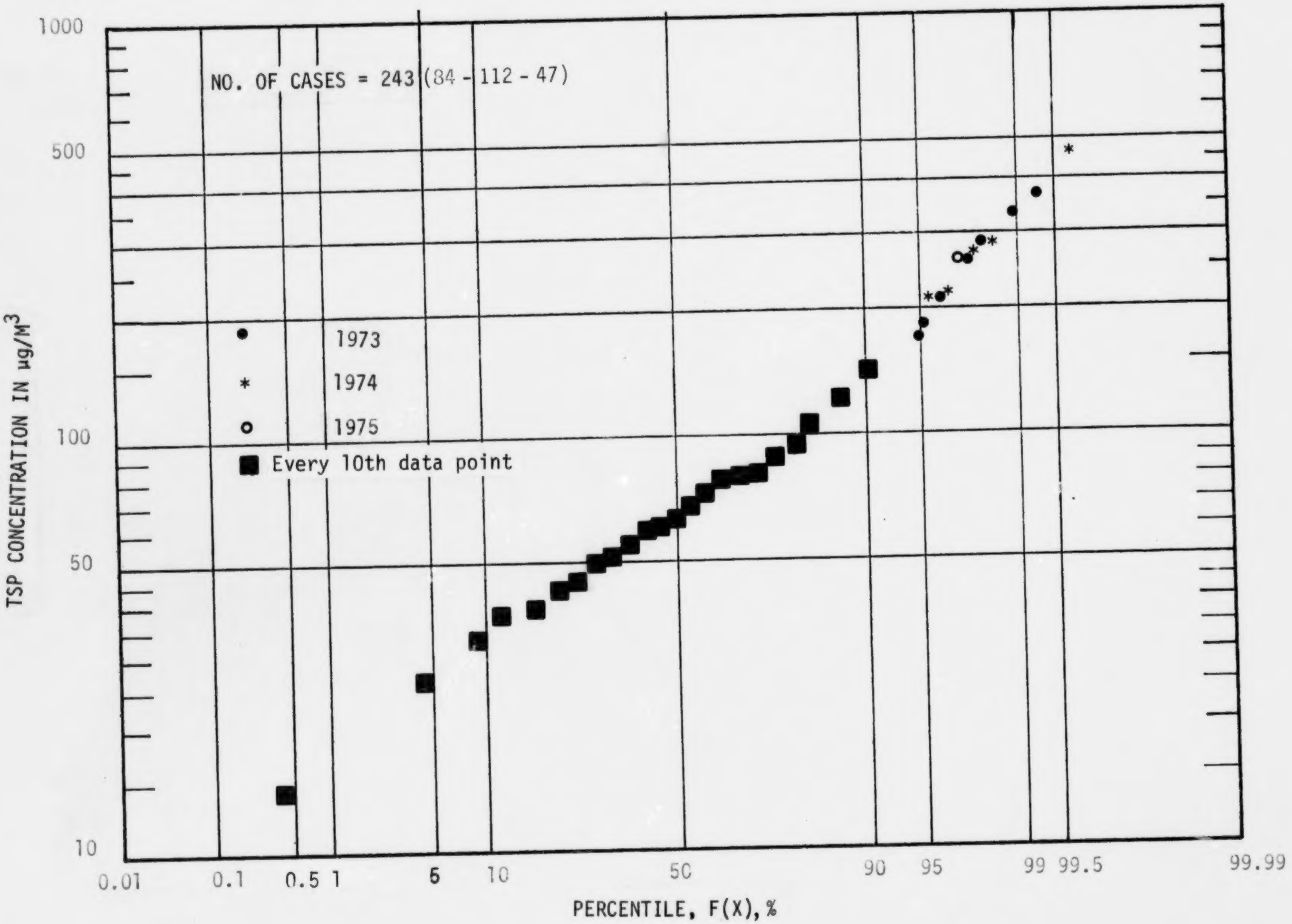


Figure A-17a. Log-Normal Distribution Plot of TSP Data at Tulsa, OK (#373000110F01)

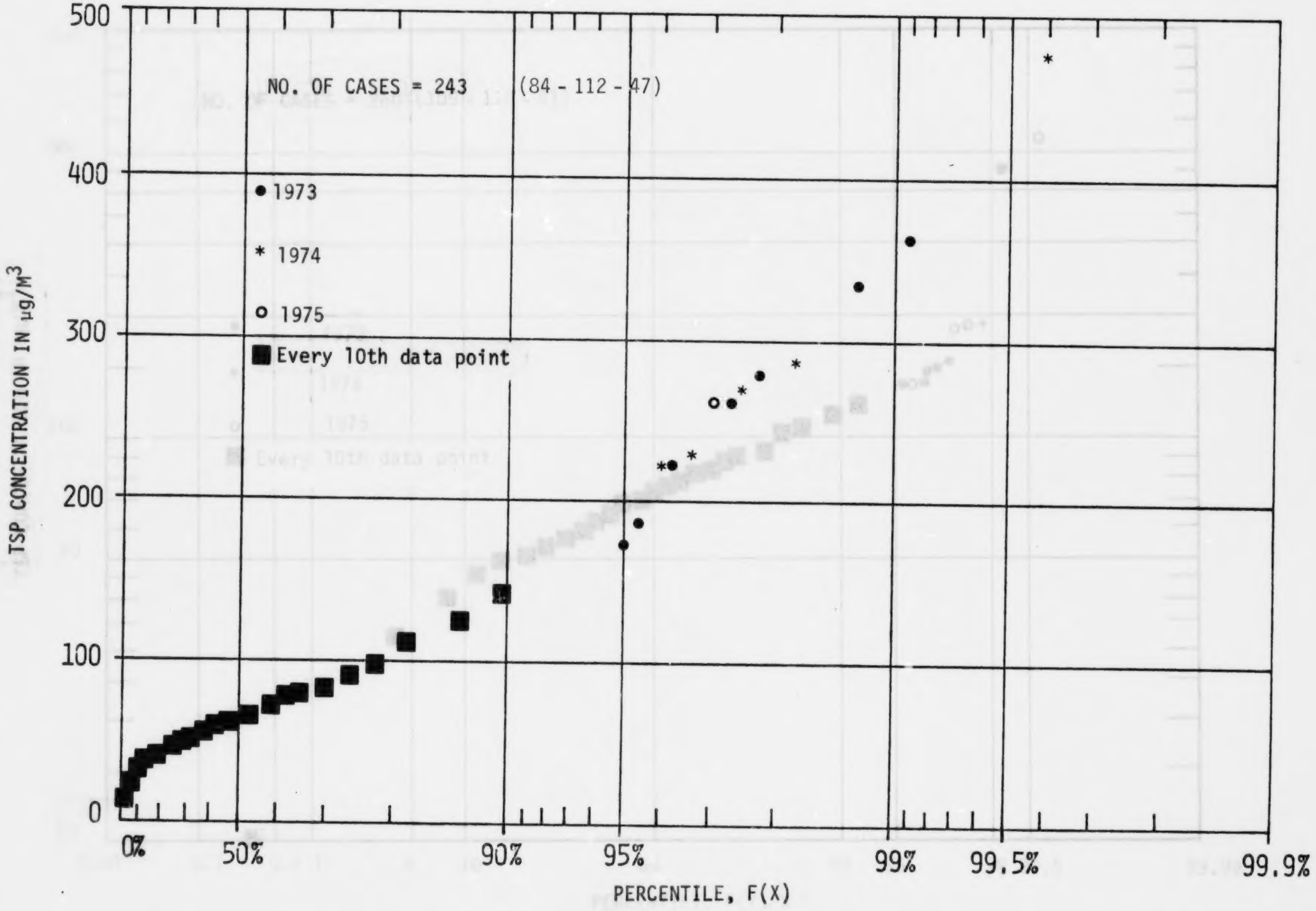


Figure A-17b. Exponential Distribution Plot of TSP Data at Tulsa, OK (#373000110F01)

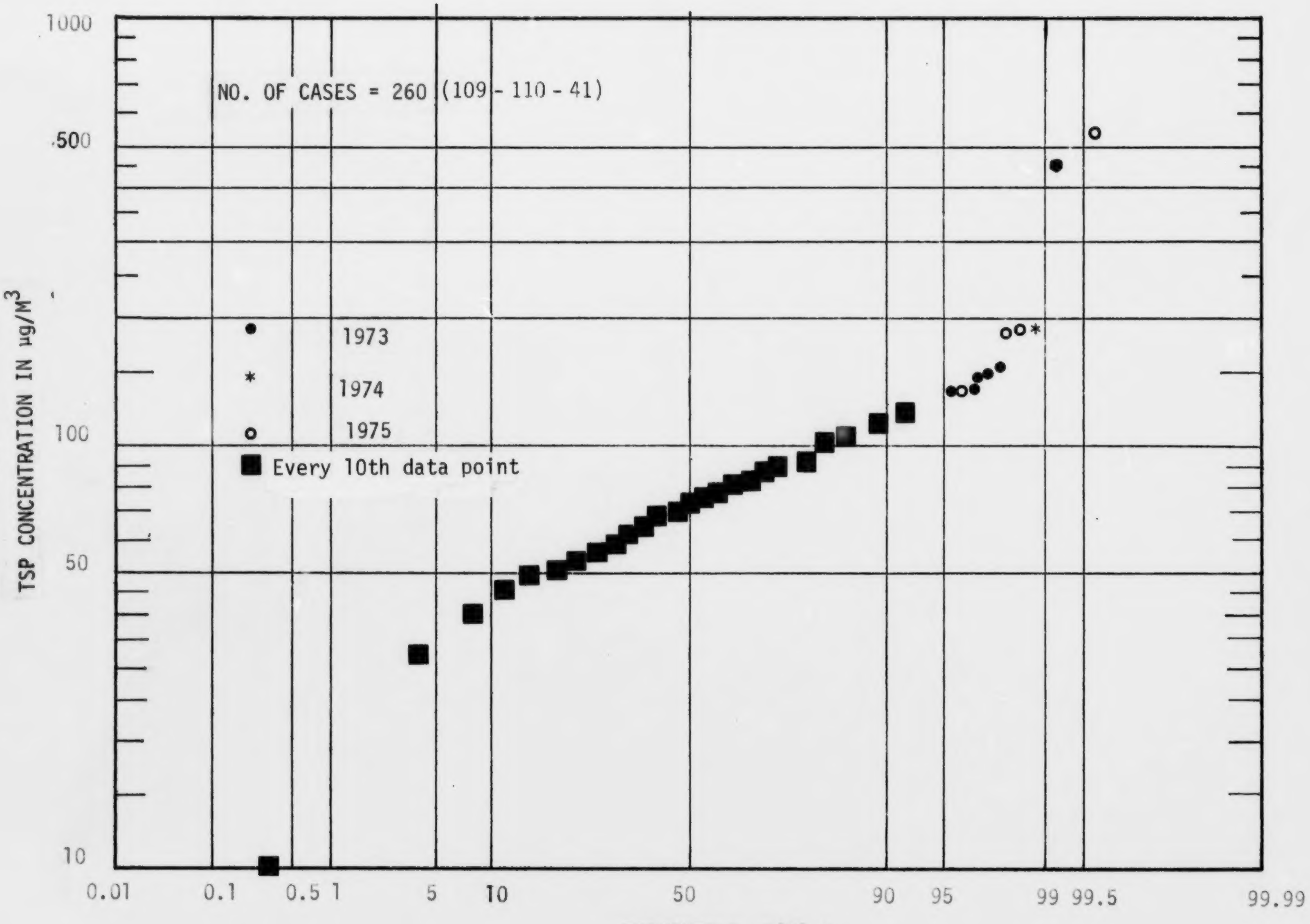


Figure A-18a. Log-Normal Distribution Plot of TSP Data at Tulsa, OK (#373000112F01)

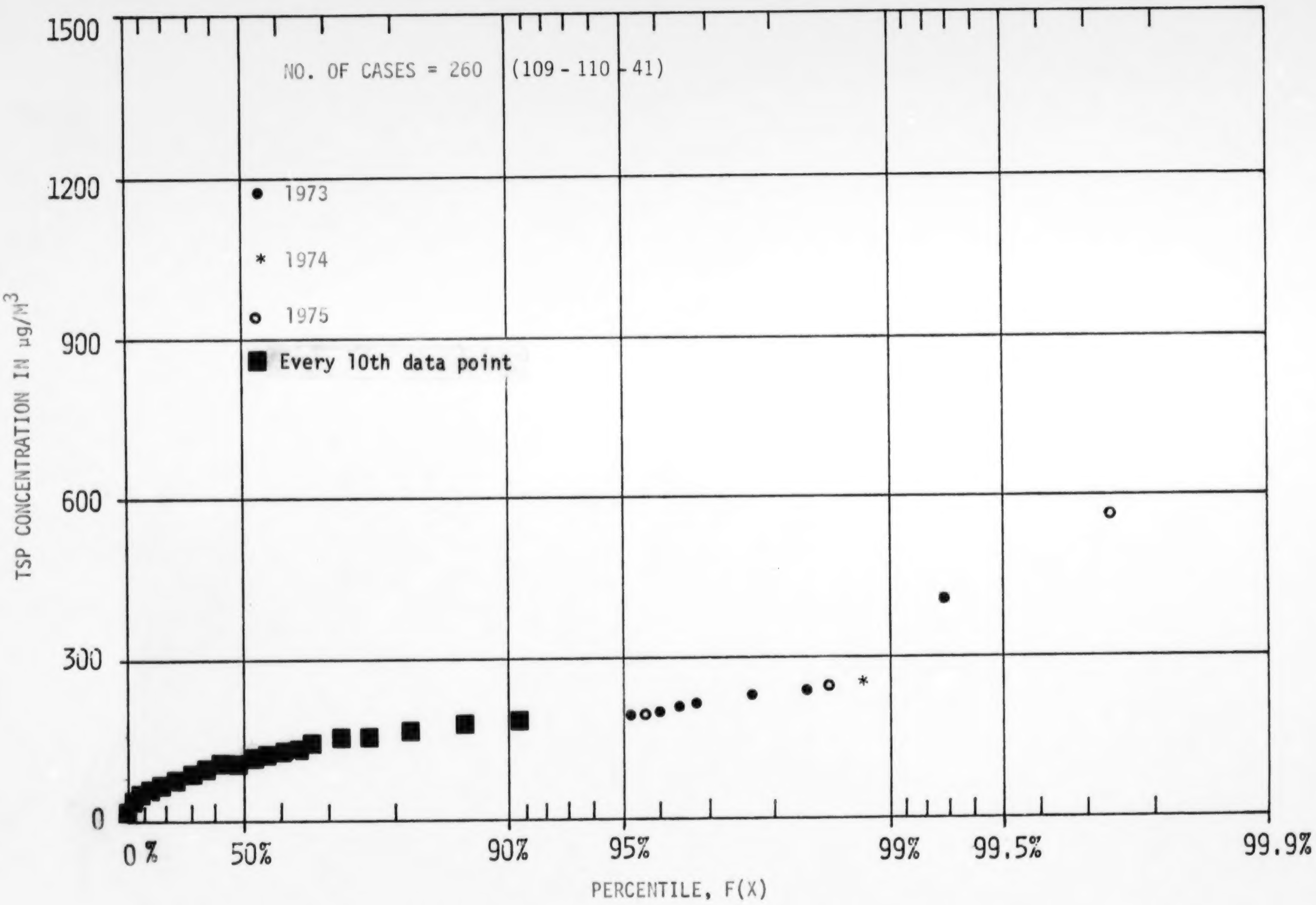


Figure A-18b. Distribution Plot of TSP Data at Tulsa, OK (#373000112F01)

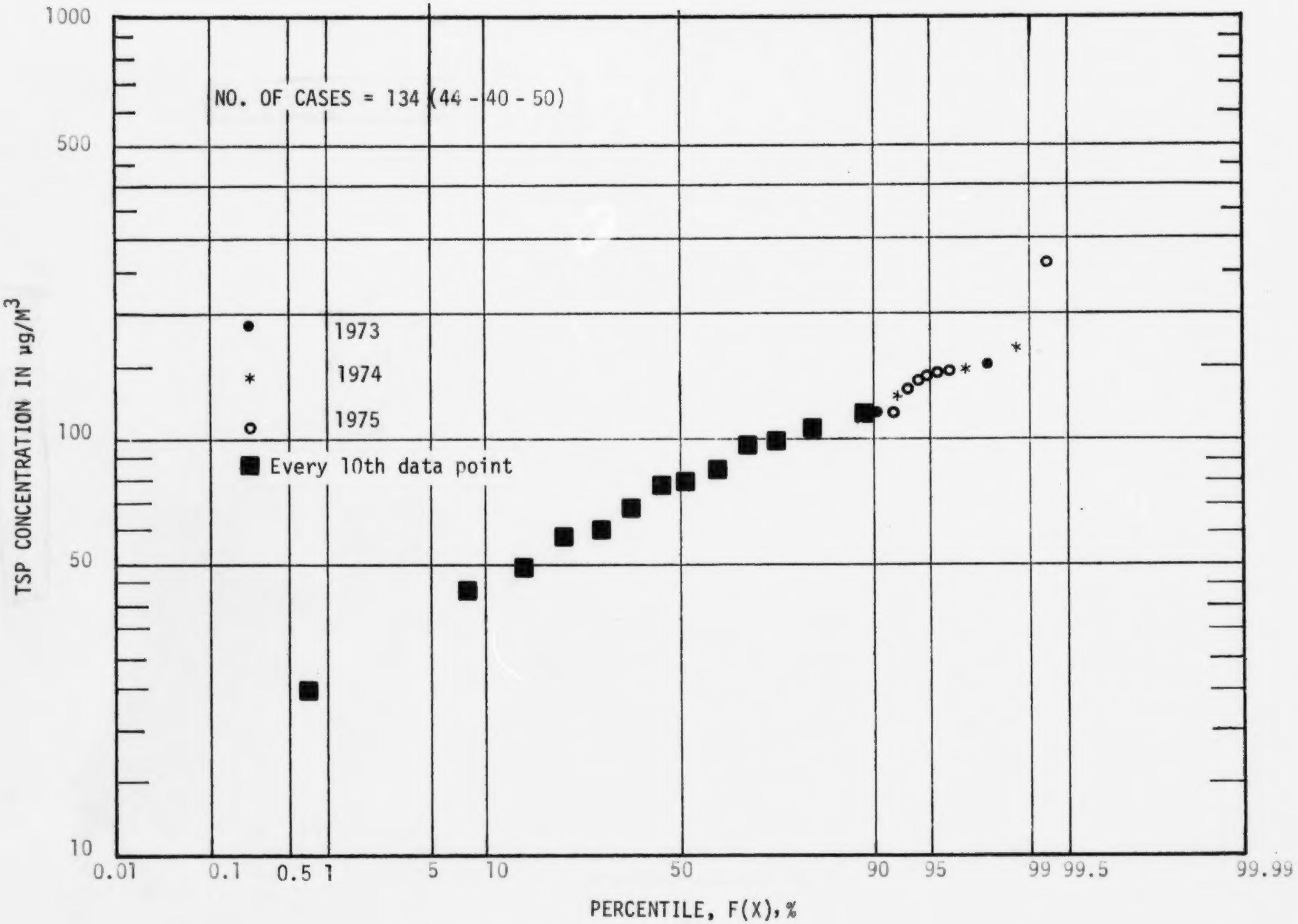


Figure A-19a. Log-Normal Distribution Plot of TSP Data at Corpus Christi, Texas (#451150001F01)

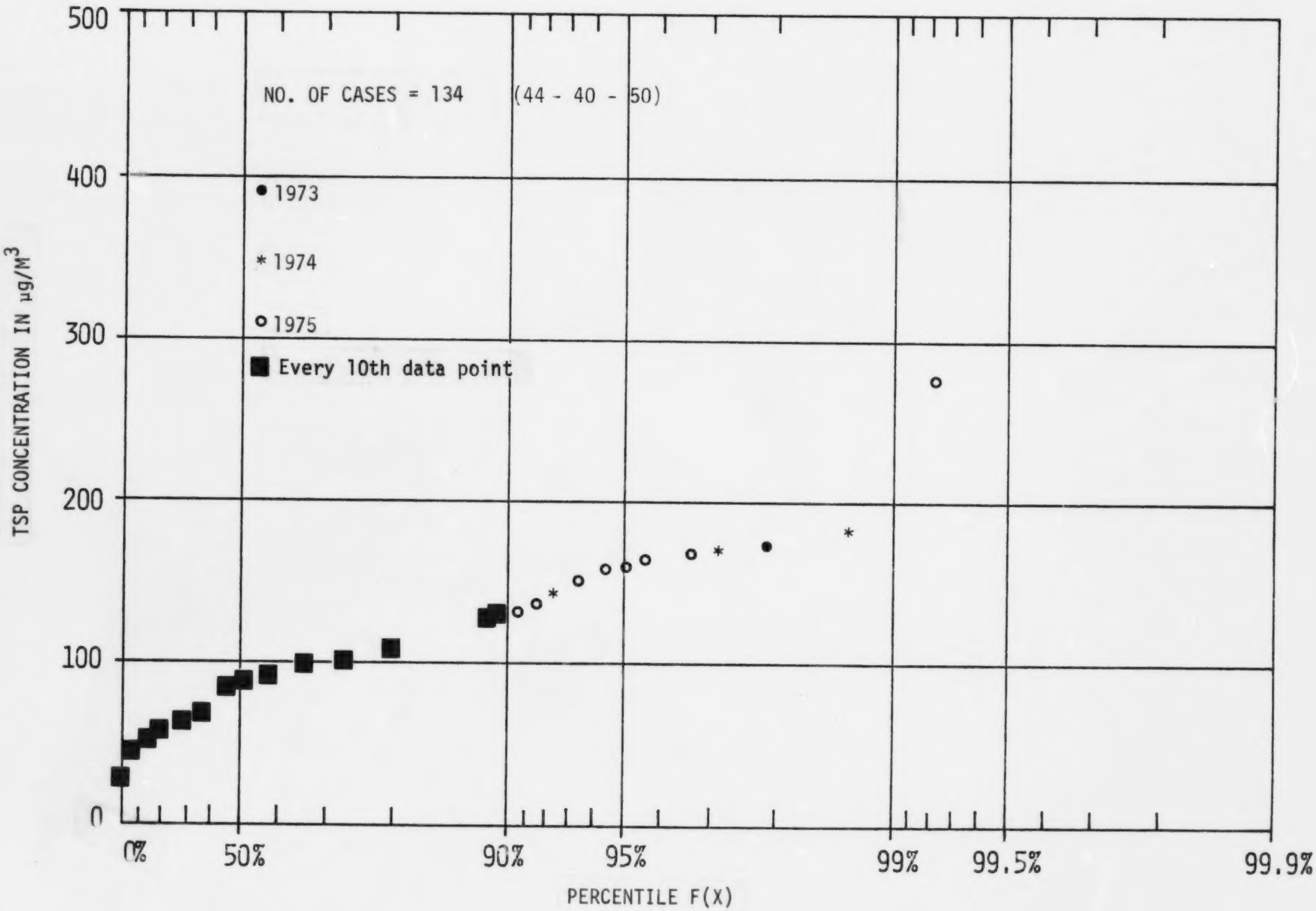


Figure A-19b. Exponential Distribution Plot of TSP Data at Corpus Christi, Texas (#451150001F01)

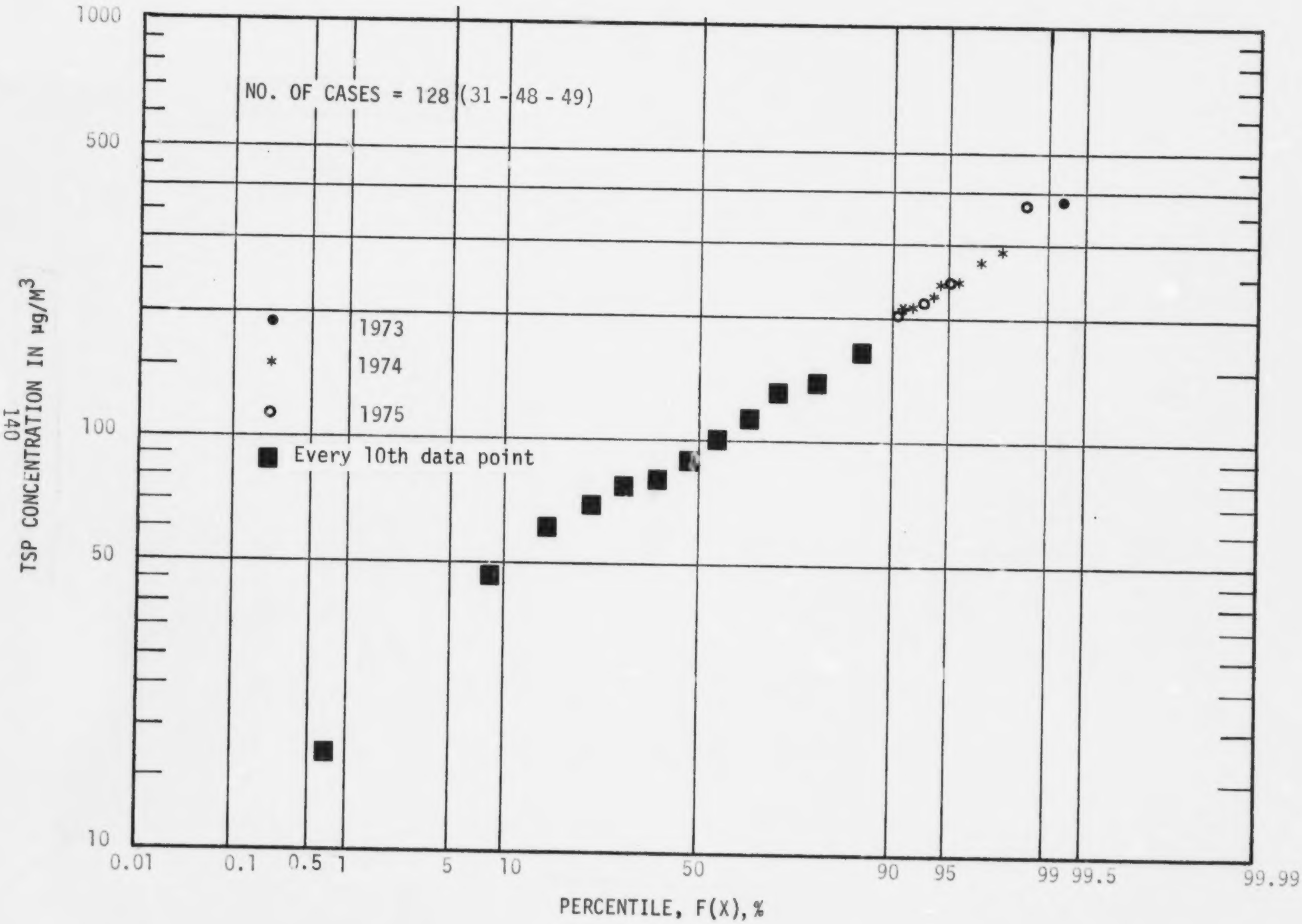


Figure A-20a. Log-Normal Distribution Plot of TSP Data at Corpus Christi, Texas (#451150004F02)

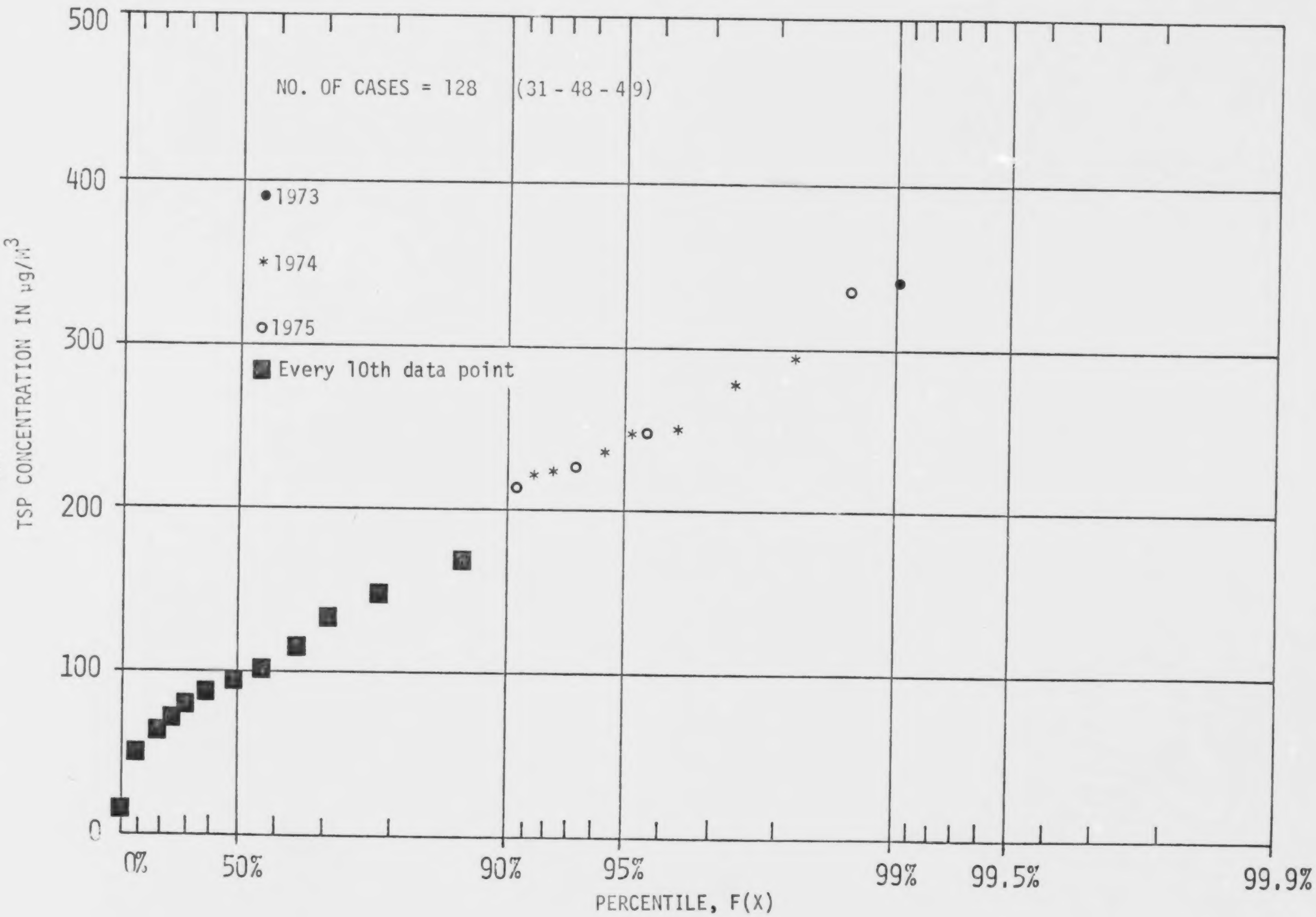


Figure A-20b. Exponential Distribution Plot of TSP Data at Corpus Christi, Texas (#451150004F02)

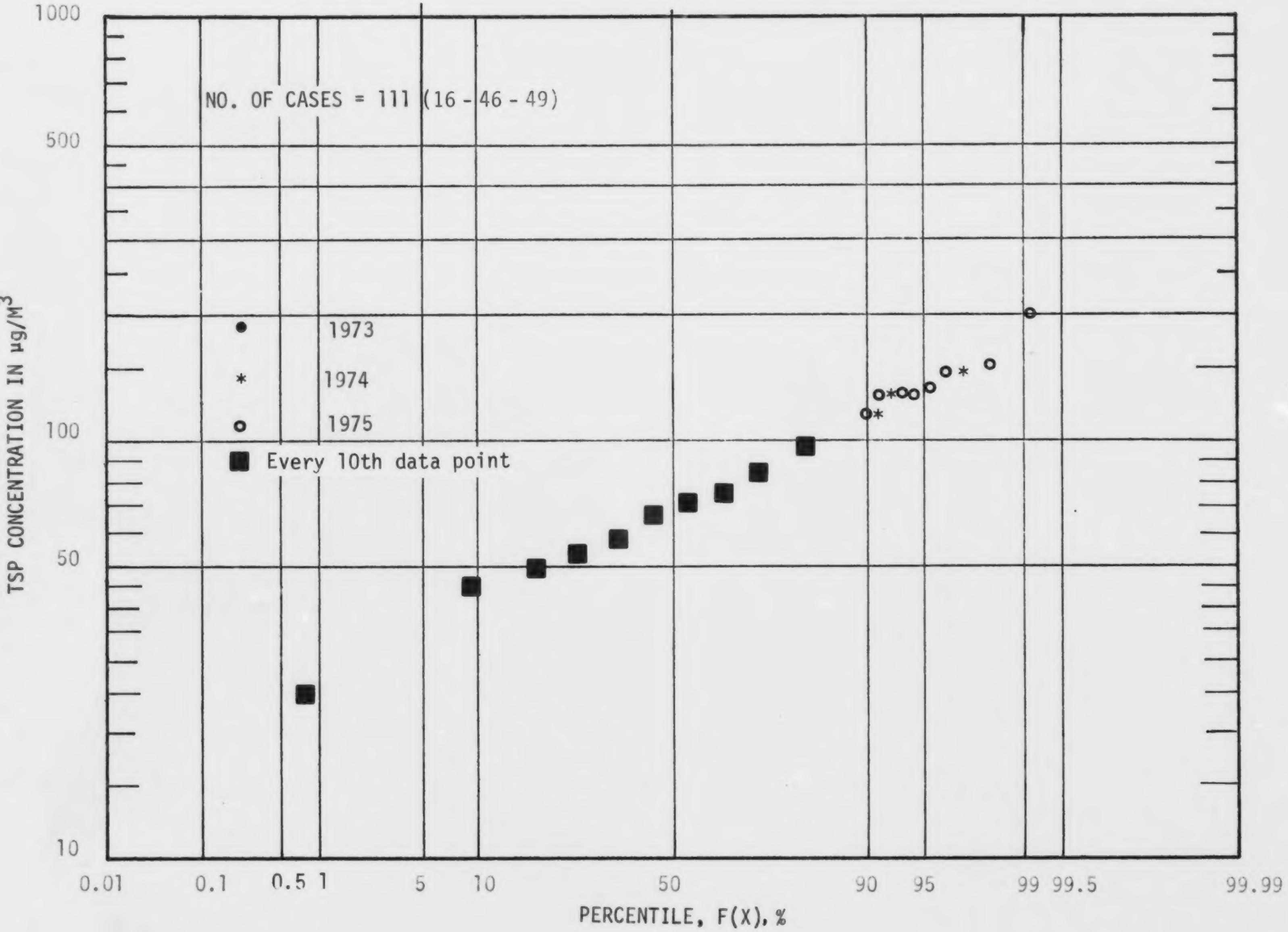


Figure A-21a. Log-Normal Distribution Plot of TSP Data at Corpus Christi, Texas (#451150019F01)

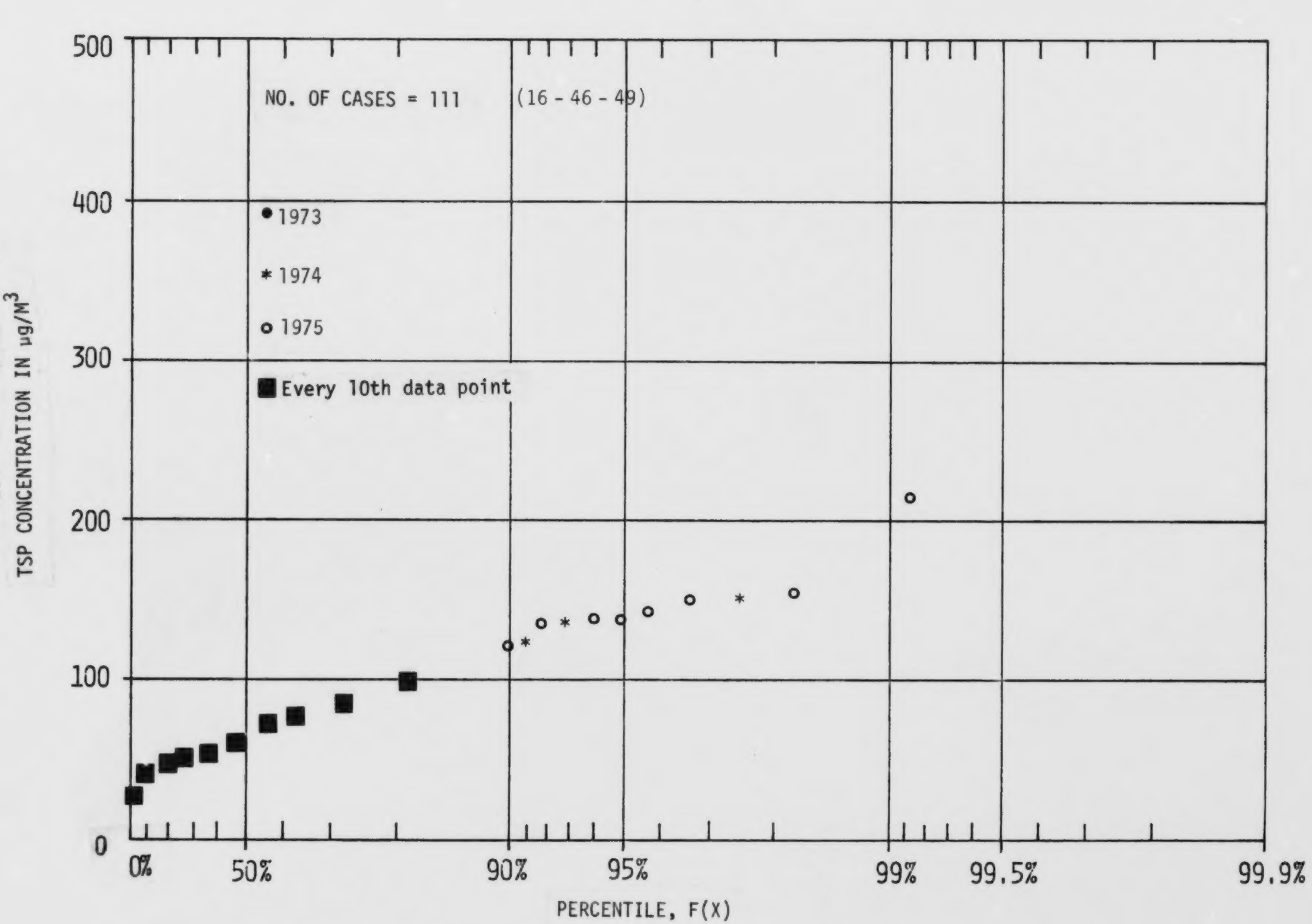


Figure A-21b. Exponential Distribution Plot of TSP Data at Corpus Christi, Texas (#451150019F01)

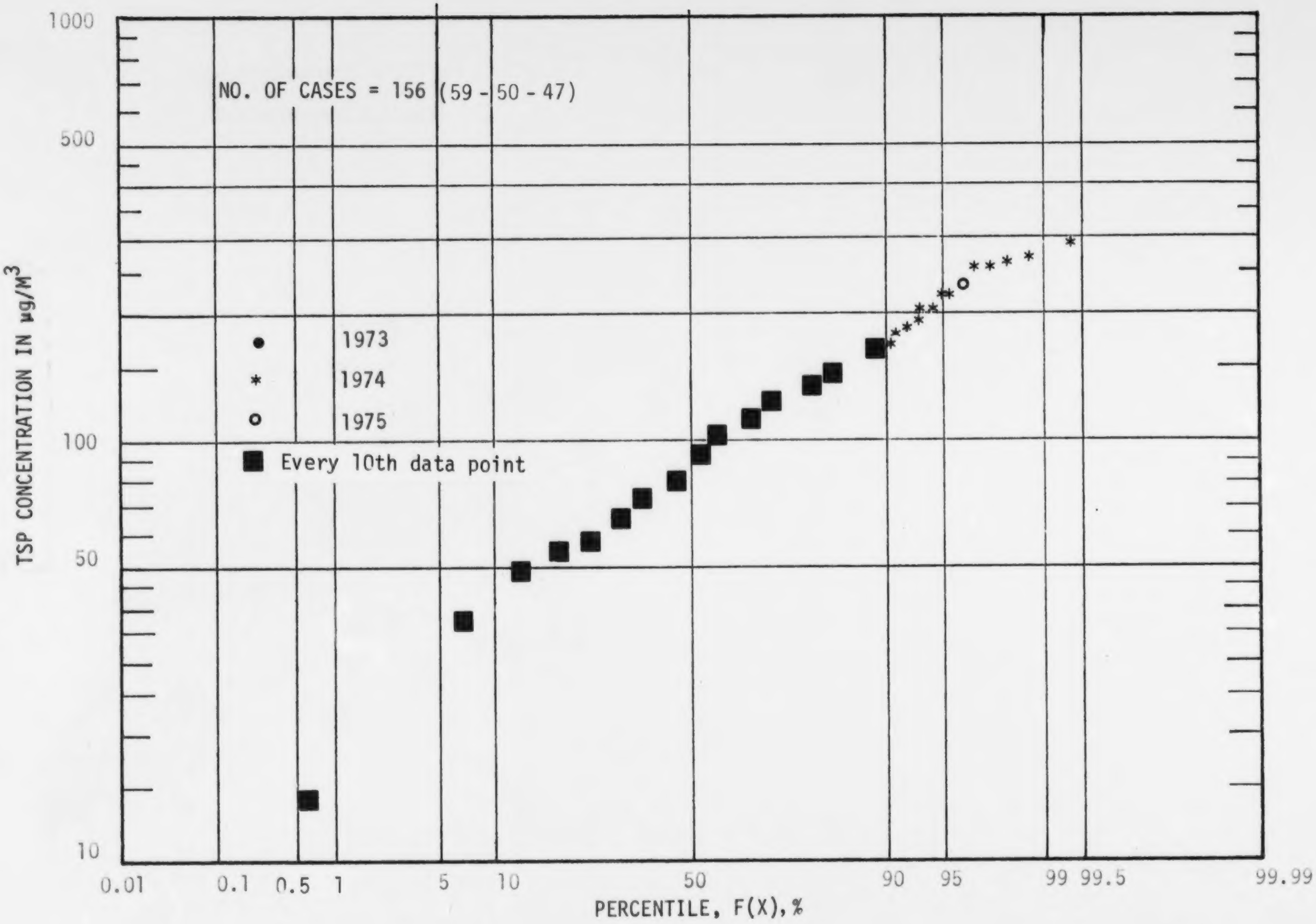


Figure A-22a. Log-Normal Distribution Plot of TSP Data at Harlingen, Texas (#452320001F01)

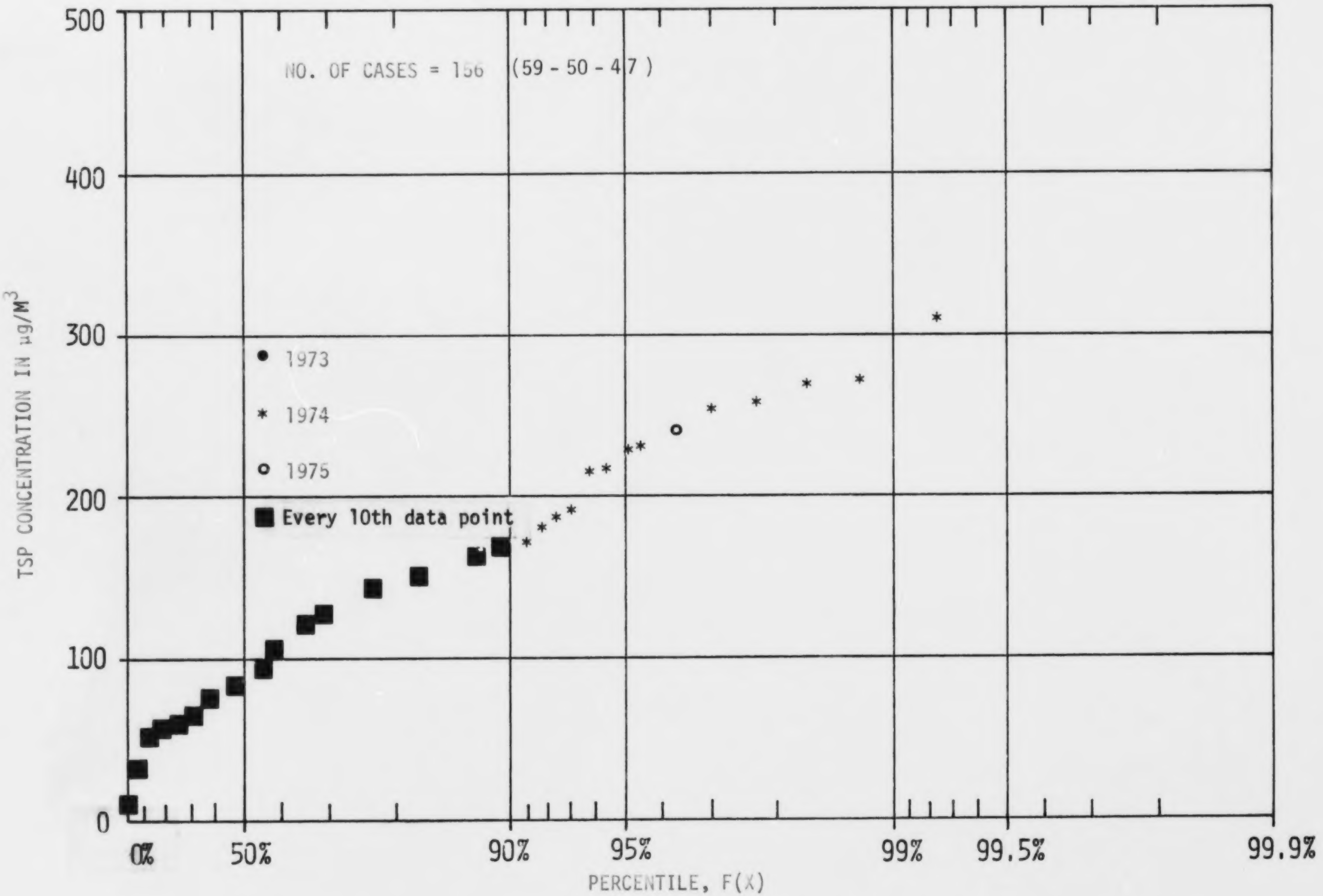


Figure A-22b. Exponential Distribution Plot of TSP Data at Harlingen, Texas (#452320001F01)

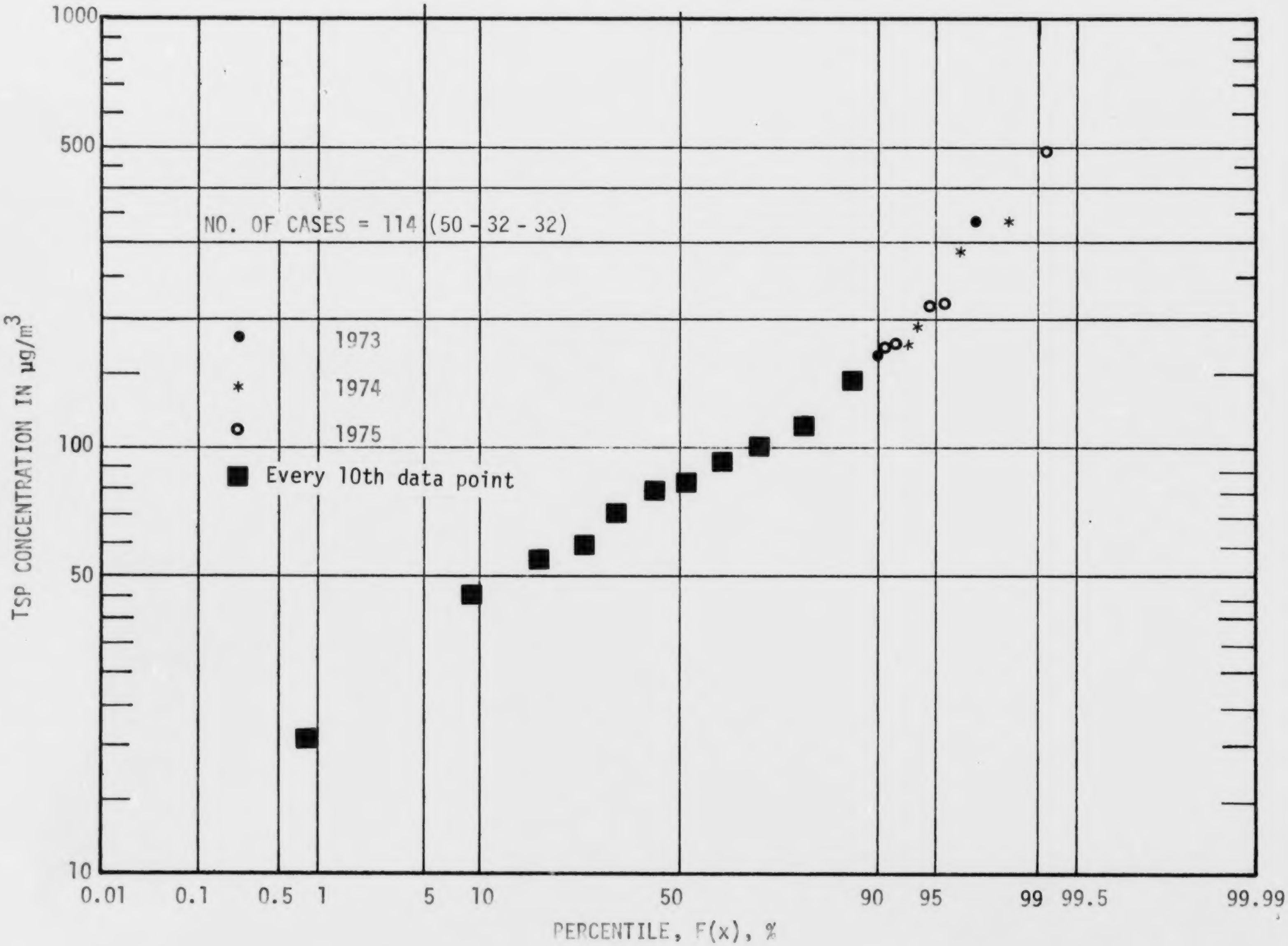


Figure A-23a. Log-normal distribution plot of TSP data at Lubbock, Texas (#453340001F01).

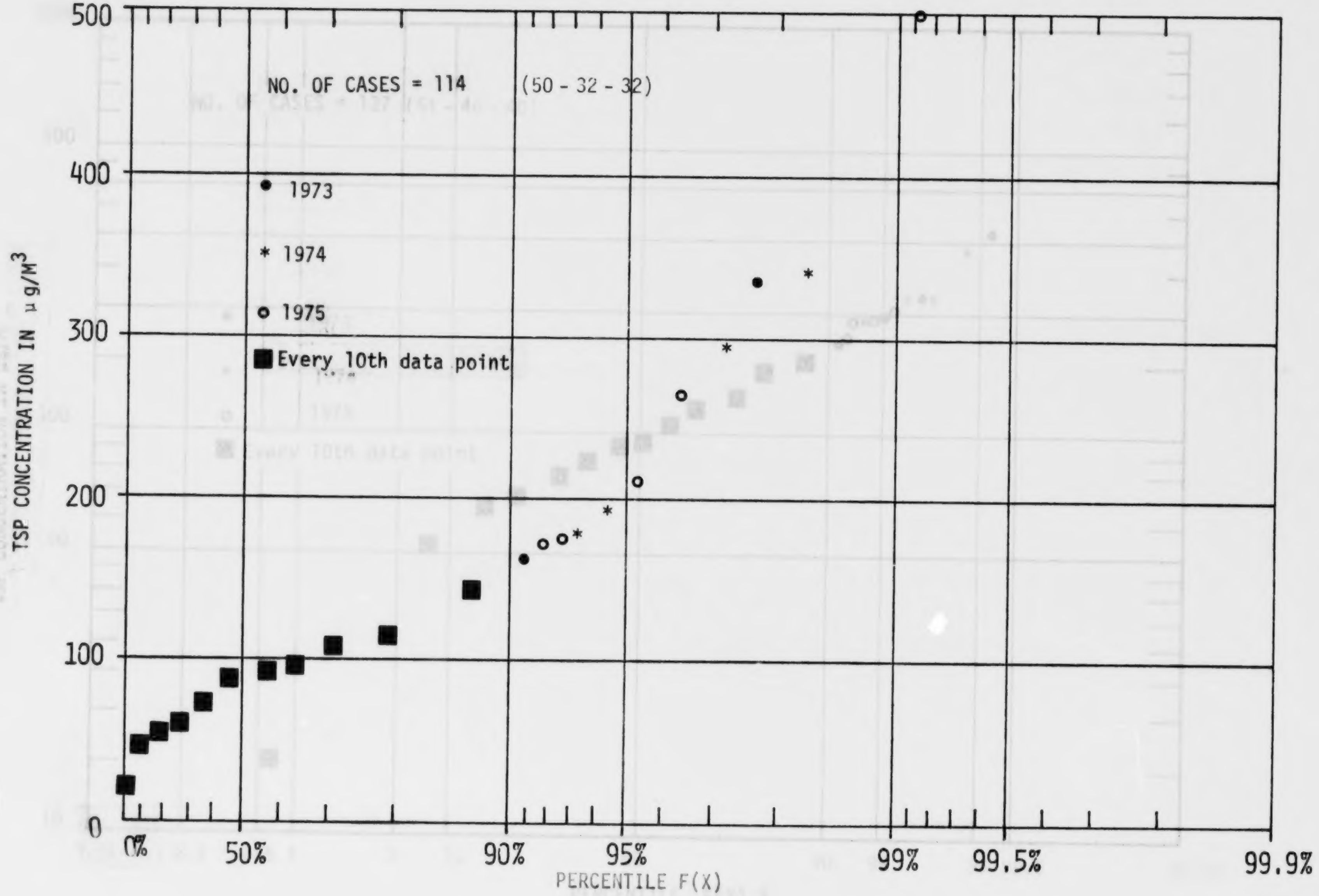


Figure A-23b. **Exponential Distribution Plot of TSP Data at Lubbock, Texas (#453340001F01)**

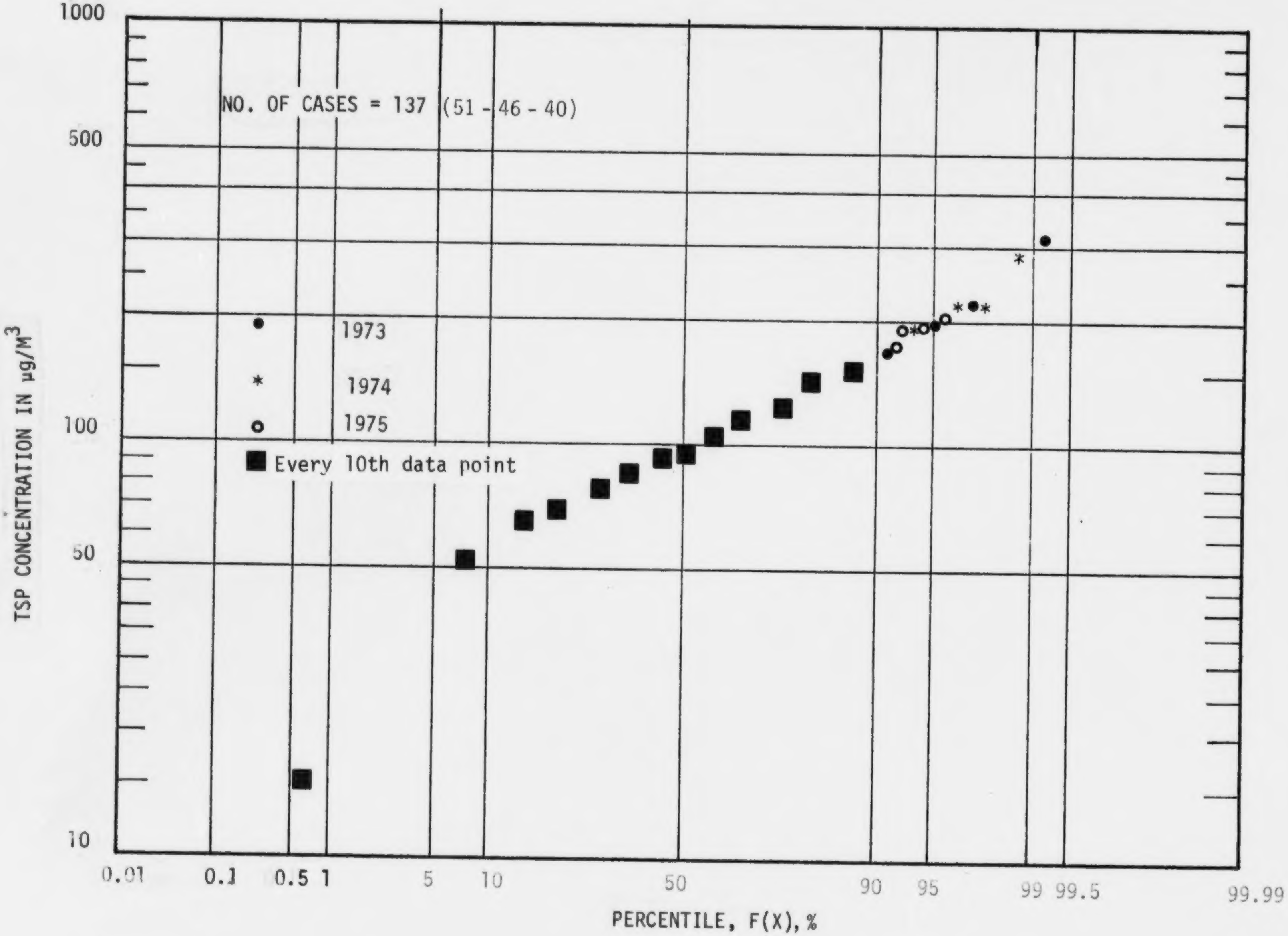


Figure A-24a. Log-Normal Distribution Plot of TSP Data at McAllen, Texas (#453390001F01)

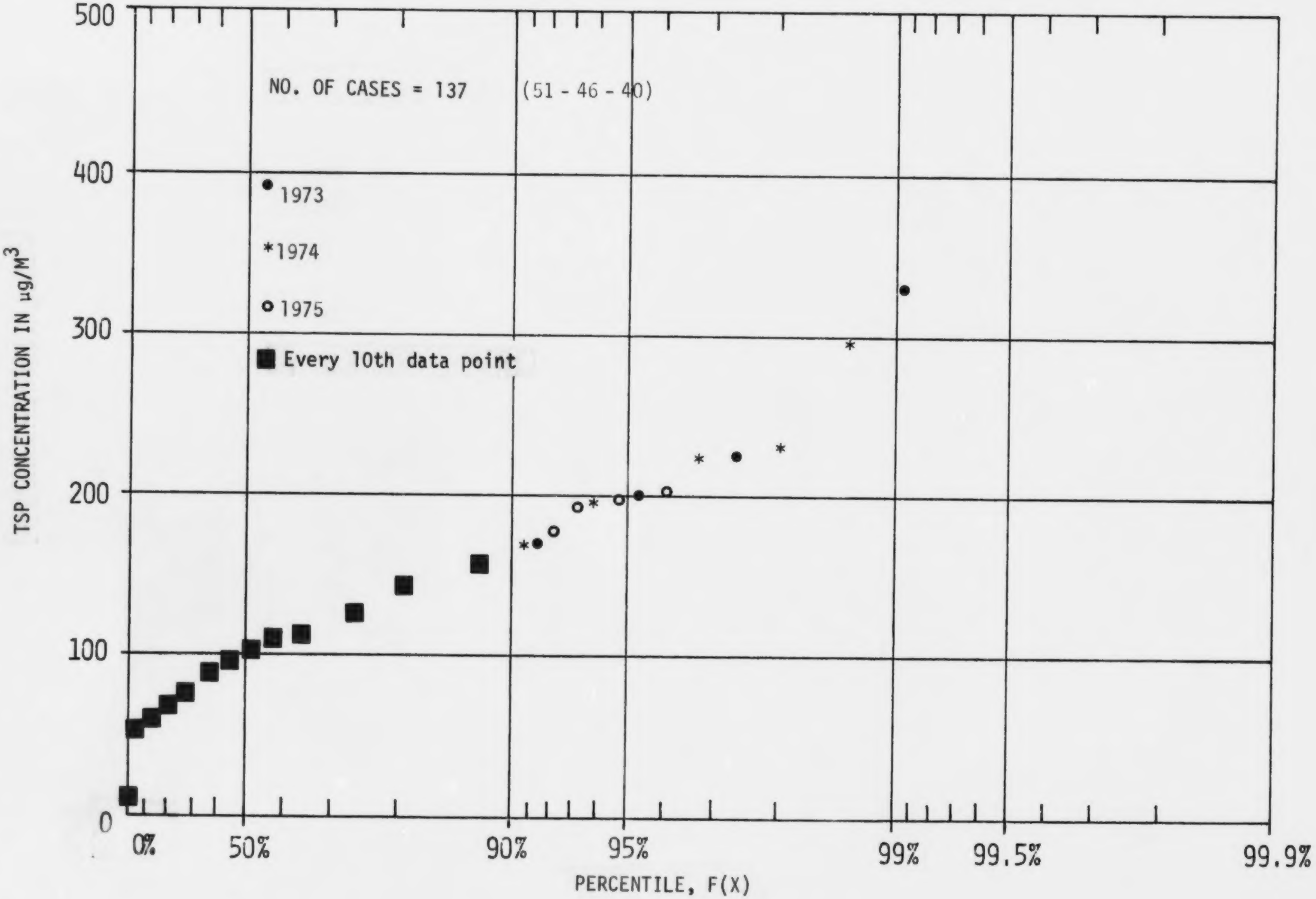


Figure A-24b. Exponential Distribution Plot of TSP Data at McAllen, Texas (#453390001F01)

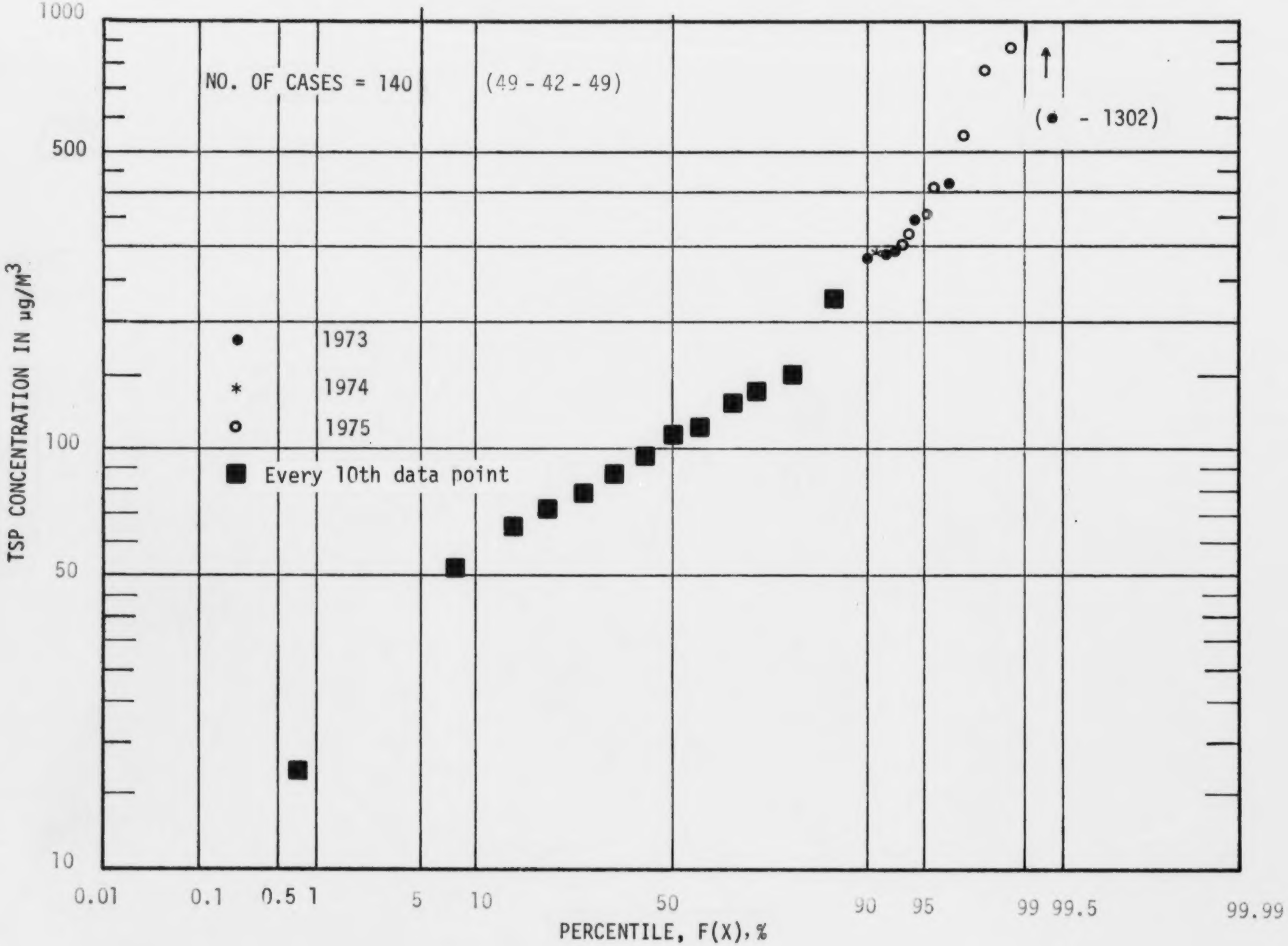


Figure A-25a. Log-Normal Distribution Plot of TSP Data at San Benito, Texas (#454600001F01)

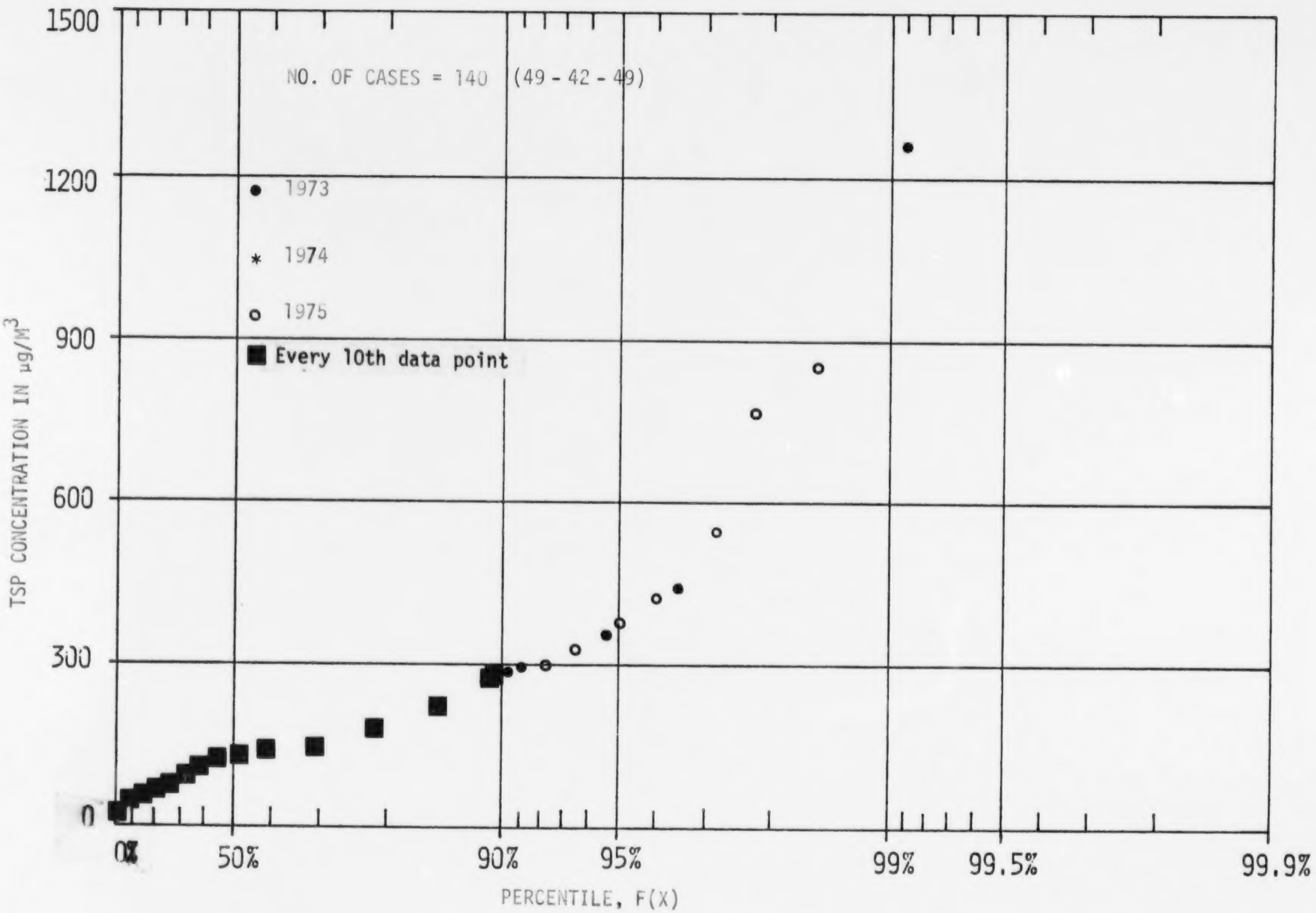


Figure A-25b. Exponential Distribution Plot of TSP Data at San Benito, Texas (#454600001F01)

APPENDIX B

AID DECISION TREES FOR THE 25 STUDY LOCATIONS

See Table 10 for
definition of variables

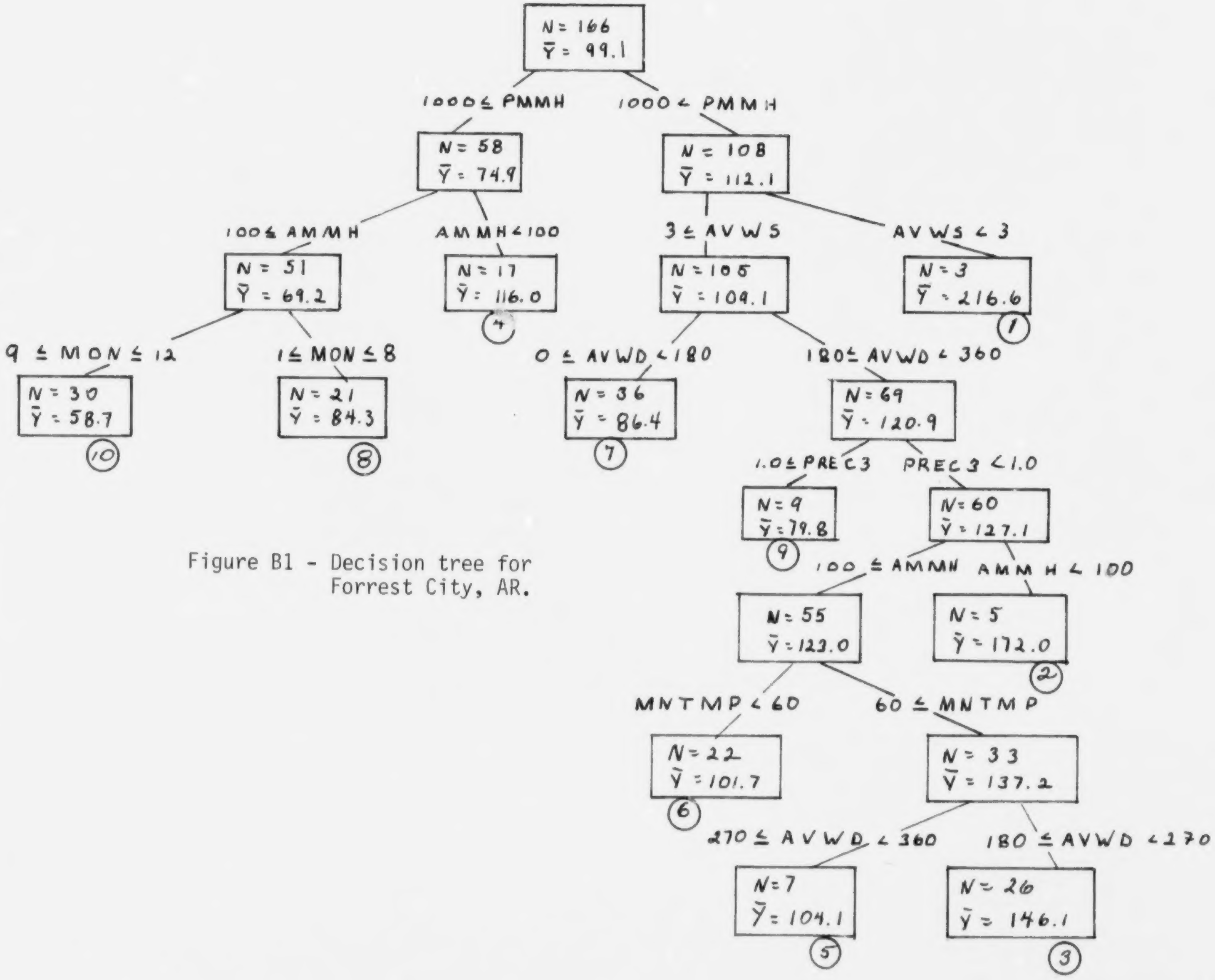

N = 166
Ȳ = 99.1
1000 ≤ PMMH
1000 < PMMH
N = 58
Ȳ = 74.9
N = 108
Ȳ = 112.1
100 ≤ AMMH
AMMH < 100
3 ≤ AVWS
AVWS < 3
N = 51
Ȳ = 69.2
N = 17
Ȳ = 116.0
④
N = 105
Ȳ = 109.1
N = 3
Ȳ = 216.6
①
9 ≤ MON ≤ 12
1 ≤ MON ≤ 8
0 ≤ AVWD < 180
180 ≤ AVWD < 360
N = 30
Ȳ = 58.7
⑩
N = 21
Ȳ = 84.3
⑧
N = 36
Ȳ = 86.4
⑦
N = 69
Ȳ = 120.9
1.0 ≤ PREC3
PREC3 < 1.0
N = 9
Ȳ = 79.8
⑨
N = 60
Ȳ = 127.1
100 ≤ AMMH
AMMH < 100
N = 55
Ȳ = 123.0
N = 5
Ȳ = 172.0
②
MNTMP < 60
60 ≤ MNTMP
N = 22
Ȳ = 101.7
⑥
N = 33
Ȳ = 137.2
270 ≤ AVWD < 360
180 ≤ AVWD < 270
N = 7
Ȳ = 104.1
⑤
N = 26
Ȳ = 146.1
③

Figure B1 - Decision tree for Forrest City, AR.

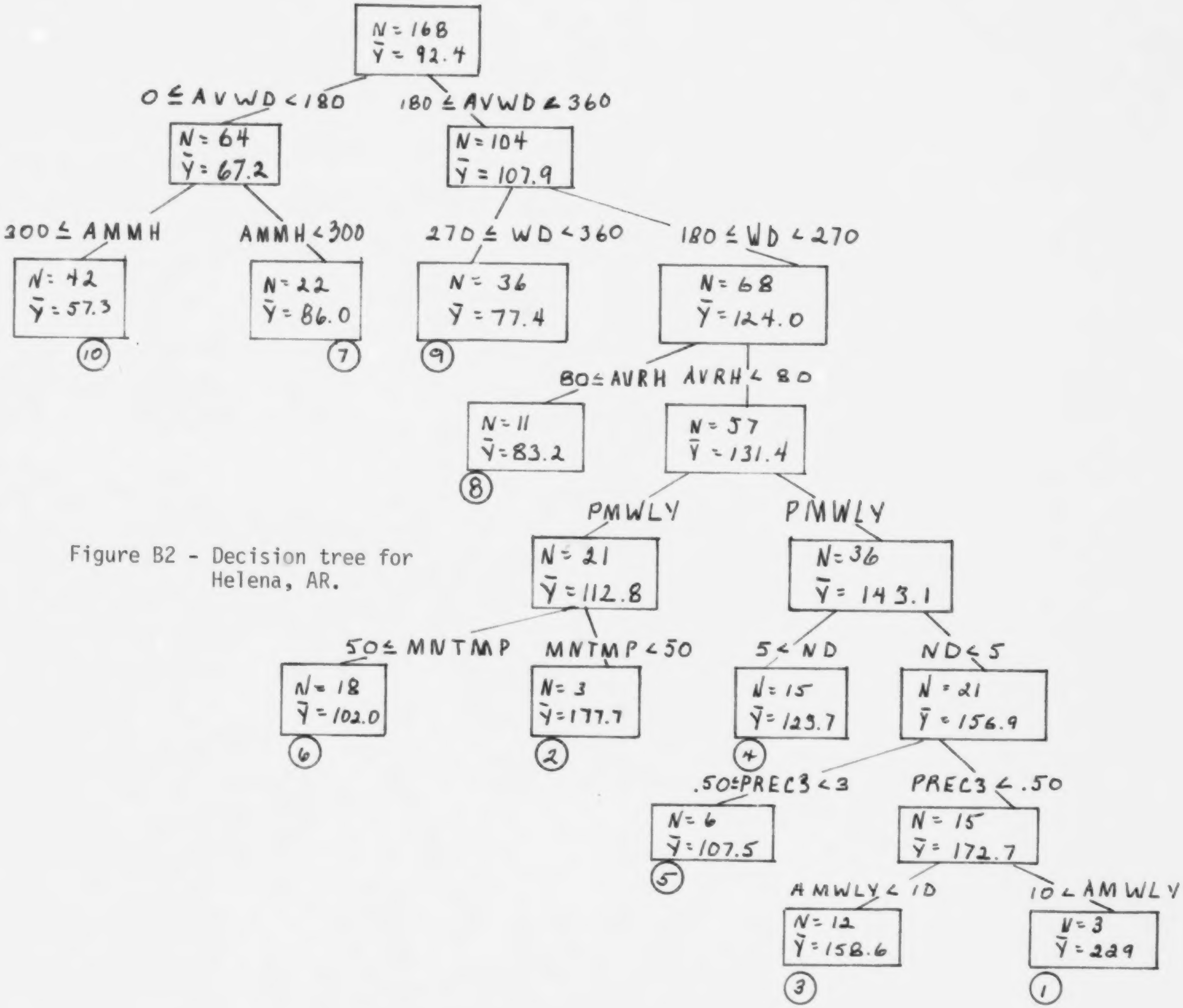

N = 168
Ȳ = 92.4
0 ≤ AVWD < 180
180 ≤ AVWD < 360
N = 64
Ȳ = 67.2
N = 104
Ȳ = 107.9
300 ≤ AMMH
AMMH < 300
270 ≤ WD < 360
180 ≤ WD < 270
N = 42
Ȳ = 57.3
10
N = 22
Ȳ = 86.0
7
N = 36
Ȳ = 77.4
9
N = 68
Ȳ = 124.0
80 ≤ AVRH
AVRH < 80
N = 11
Ȳ = 83.2
8
N = 57
Ȳ = 131.4
PMWLY
PMWLY
N = 21
Ȳ = 112.8
N = 36
Ȳ = 143.1
50 ≤ MNTMP
MNTMP < 50
5 < ND
ND < 5
N = 18
Ȳ = 102.0
6
N = 3
Ȳ = 177.7
2
N = 15
Ȳ = 123.7
4
N = 21
Ȳ = 156.9
.50 ≤ PREC3 < 3
PREC3 < .50
N = 6
Ȳ = 107.5
5
N = 15
Ȳ = 172.7
AMWLY < 10
10 < AMWLY
N = 12
Ȳ = 158.6
3
N = 3
Ȳ = 229
1

Figure B2 - Decision tree for Helena, AR.

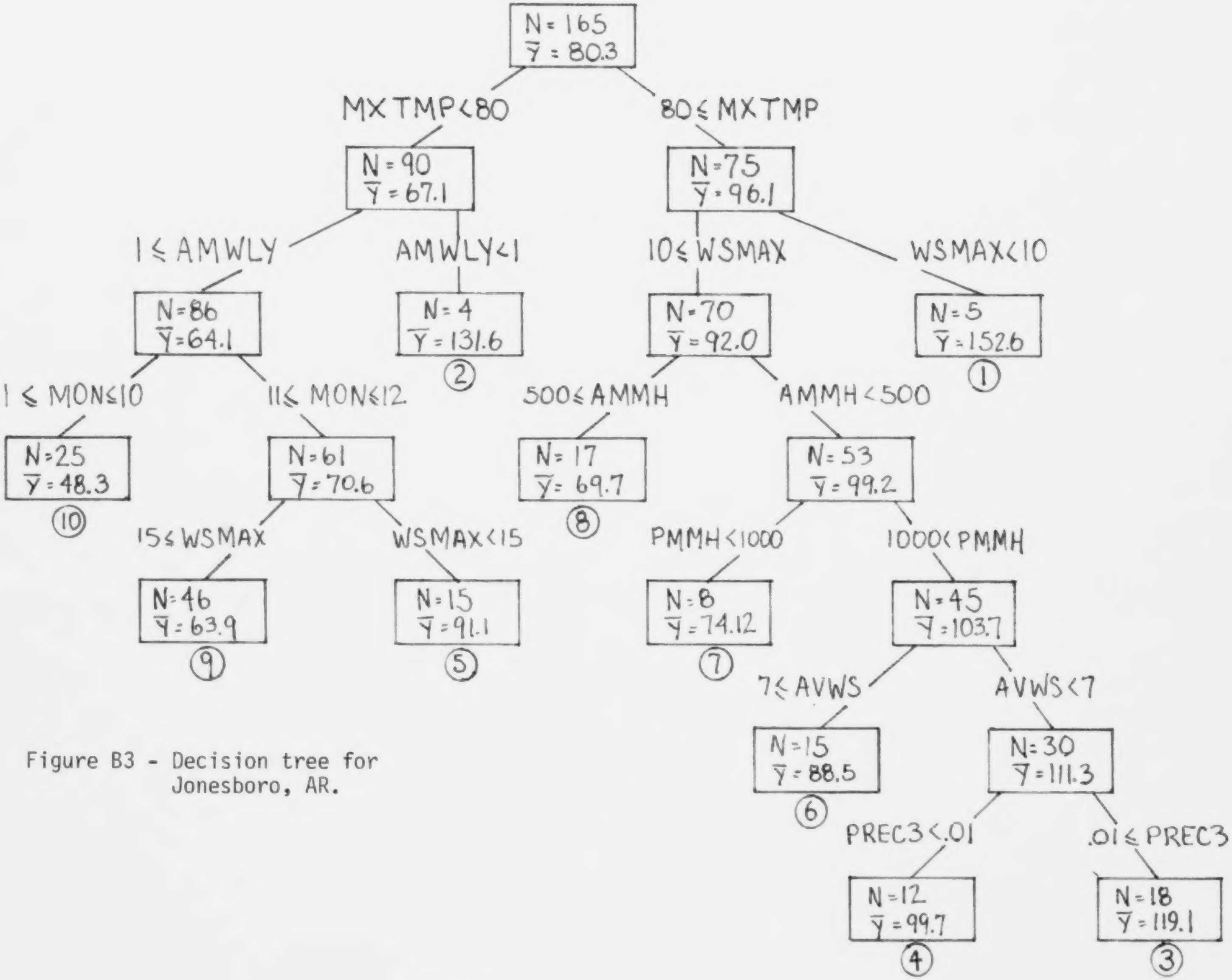

N = 165
$\bar{Y} = 80.3$
MXTMP<80
80≤MXTMP
N = 90
$\bar{Y} = 67.1$
N = 75
$\bar{Y} = 96.1$
1≤AMWLY
AMWLY<1
10≤WSMAX
WSMAX<10
N = 86
$\bar{Y} = 64.1$
N = 4
$\bar{Y} = 131.6$
②
N = 70
$\bar{Y} = 92.0$
N = 5
$\bar{Y} = 152.6$
①
1≤MON≤10
11≤MON≤12
500≤AMMH
AMMH<500
N = 25
$\bar{Y} = 48.3$
⑩
N = 61
$\bar{Y} = 70.6$
N = 17
$\bar{Y} = 69.7$
⑧
N = 53
$\bar{Y} = 99.2$
15≤WSMAX
WSMAX<15
PMMH<1000
1000<PMMH
N = 46
$\bar{Y} = 63.9$
⑨
N = 15
$\bar{Y} = 91.1$
⑤
N = 8
$\bar{Y} = 74.12$
⑦
N = 45
$\bar{Y} = 103.7$
7≤AVWS
AVWS<7
N = 15
$\bar{Y} = 88.5$
⑥
N = 30
$\bar{Y} = 111.3$
PREC3<.01
.01≤PREC3
N = 12
$\bar{Y} = 99.7$
④
N = 18
$\bar{Y} = 119.1$
③

Figure B3 - Decision tree for Jonesboro, AR.

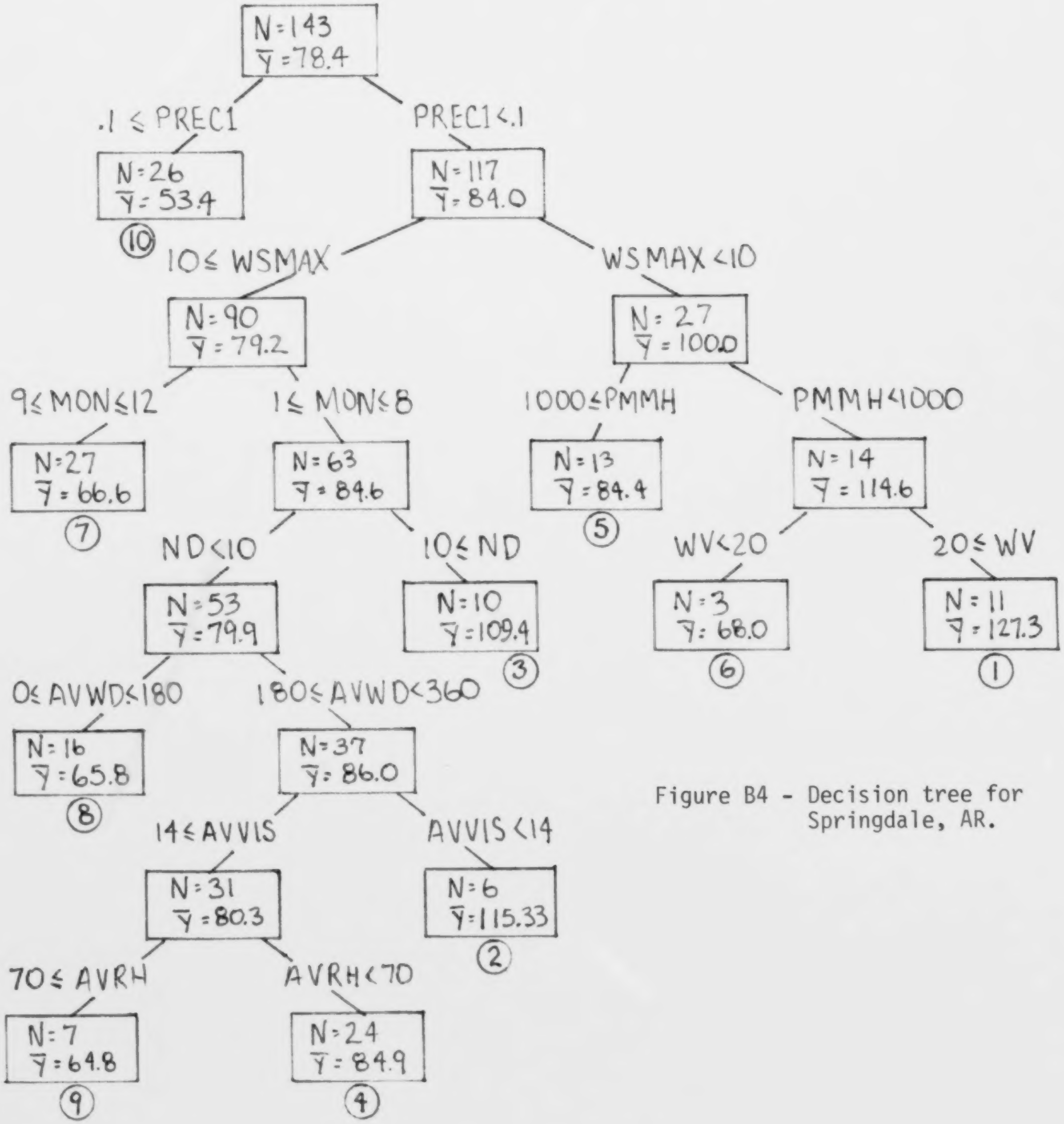

N=143
$\overline{Y}$=78.4
.1 ≤ PREC1
PREC1<.1
N=26
$\overline{Y}$=53.4
⑩
N=117
$\overline{Y}$=84.0
10≤ WSMAX
WSMAX<10
N=90
$\overline{Y}$=79.2
N=27
$\overline{Y}$=100.0
9≤ MON≤12
1≤ MON≤8
1000≤PMMH
PMMH<1000
N=27
$\overline{Y}$=66.6
⑦
N=63
$\overline{Y}$=84.6
N=13
$\overline{Y}$=84.4
⑤
N=14
$\overline{Y}$=114.6
ND<10
10≤ND
WV<20
20≤WV
N=53
$\overline{Y}$=79.9
N=10
$\overline{Y}$=109.4
③
N=3
$\overline{Y}$=68.0
⑥
N=11
$\overline{Y}$=127.3
①
0≤AVWD<180
180≤AVWD<360
N=16
$\overline{Y}$=65.8
⑧
N=37
$\overline{Y}$=86.0
14≤AVVIS
AVVIS<14
N=31
$\overline{Y}$=80.3
N=6
$\overline{Y}$=115.33
②
70≤AVRH
AVRH<70
N=7
$\overline{Y}$=64.8
⑨
N=24
$\overline{Y}$=84.9
④

Figure B4 - Decision tree for Springdale, AR.

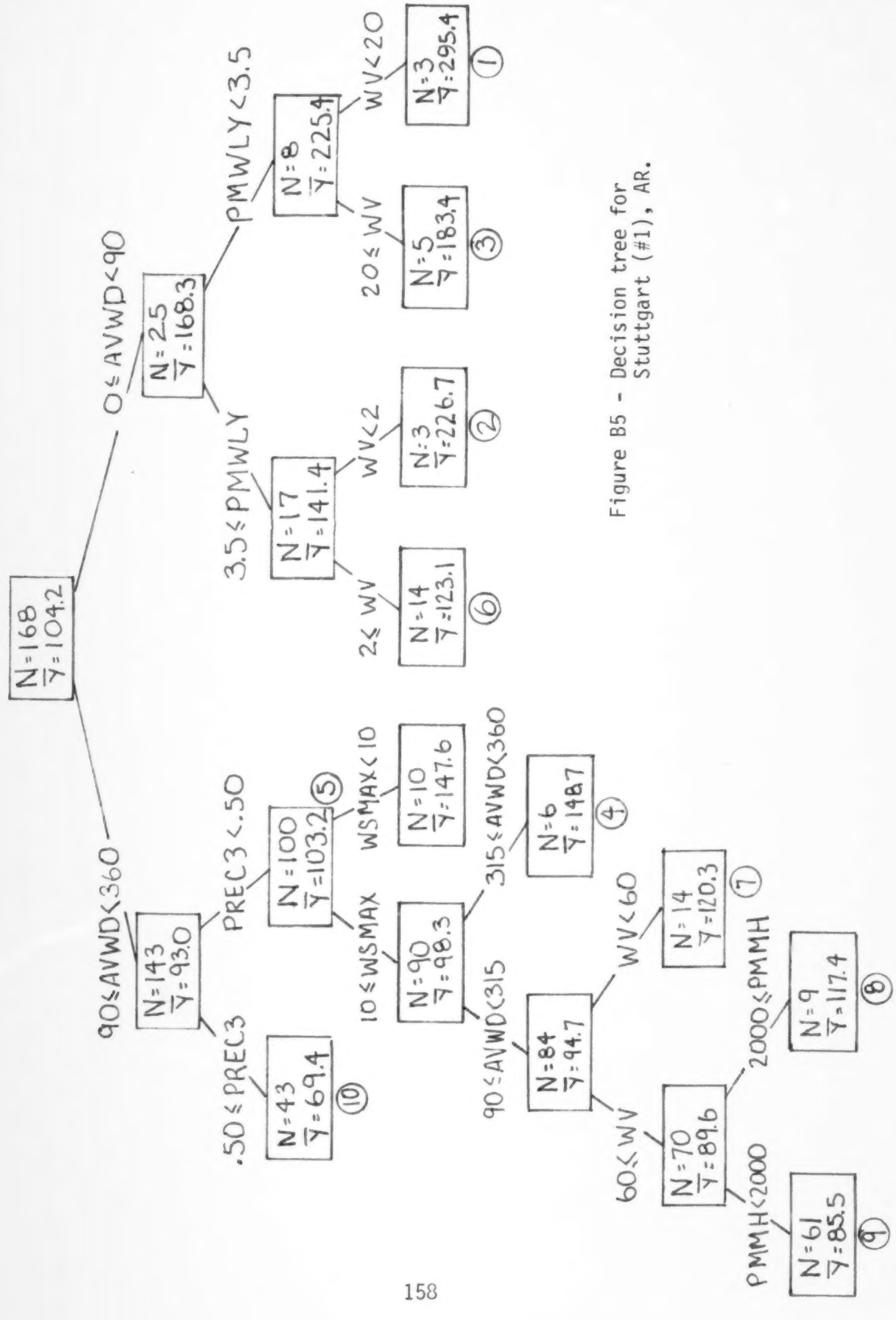

N=168
Ȳ:104.2
90≤AVWD<360
0≤AVWD<90
N=143
Ȳ:93.0
N=25
Ȳ:168.3
.50≤PREC3
PREC3<.50
3.5≤PMWLY
PMWLY<3.5
N=43
Ȳ=69.4
⑩
N=100
Ȳ=103.2
⑤
N=17
Ȳ:141.4
N=8
Ȳ:225.4
10≤WSMAX
WSMAX<10
2≤WV
WV<2
20≤WV
WV<20
N=90
Ȳ:98.3
N=10
Ȳ:147.6
N=14
Ȳ:123.1
⑥
N=3
Ȳ:226.7
②
N=5
Ȳ:183.4
③
N=3
Ȳ:295.4
①
90≤AVWD<315
315≤AVWD<360
N=84
Ȳ:94.7
N=6
Ȳ:148.7
④
60≤WV
WV<60
N=70
Ȳ:89.6
N=14
Ȳ:120.3
⑦
PMMH<2000
2000≤PMMH
N=61
Ȳ:85.5
⑨
N=9
Ȳ:117.4
⑧

Figure B5 - Decision tree for Stuttgart (#1), AR.

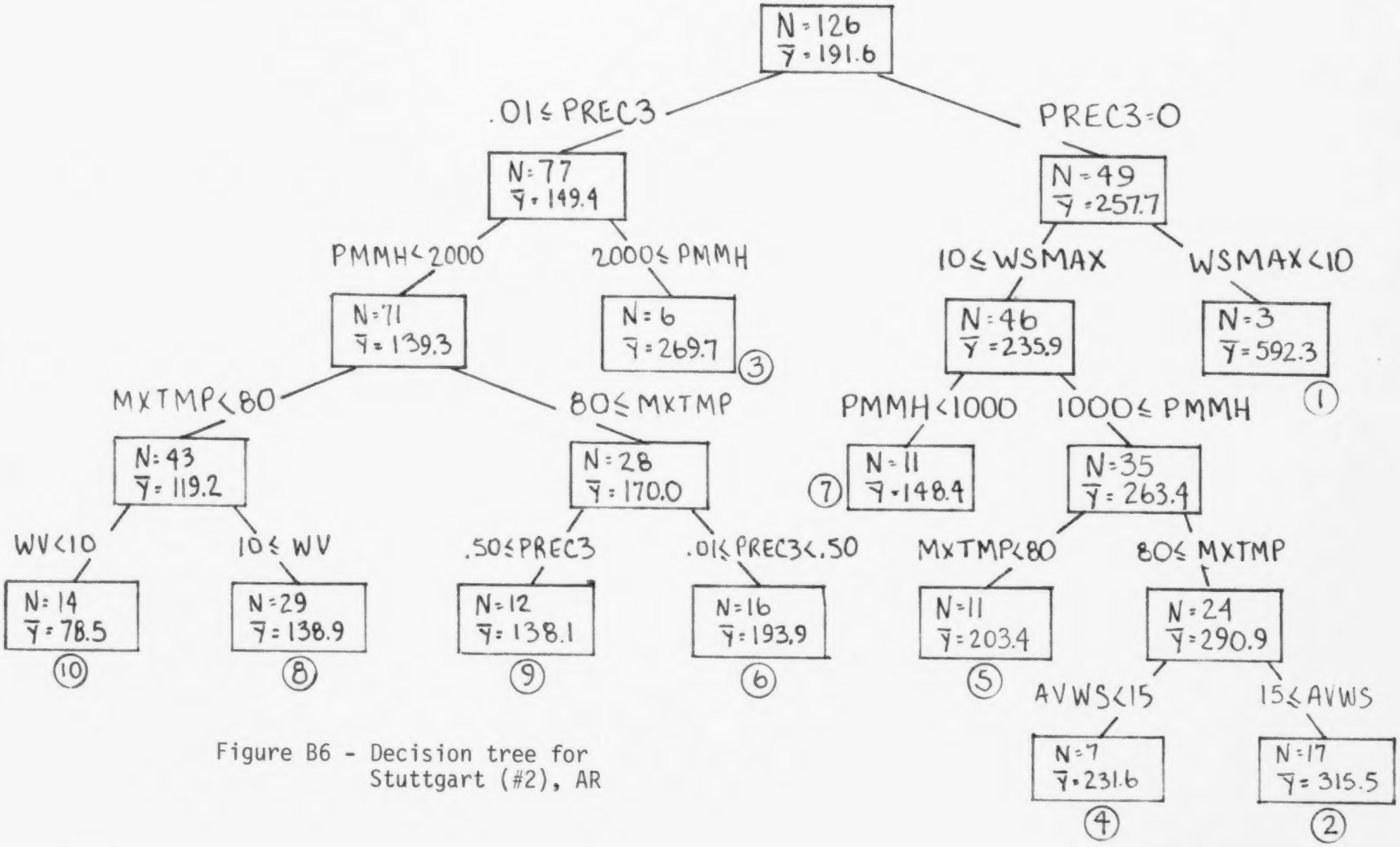

N = 126
$\bar{y}$ = 191.6
.01 ≤ PREC3
PREC3 = 0
N: 77
$\bar{y}$ = 149.4
N = 49
$\bar{y}$ = 257.7
PMMH < 2000
2000 ≤ PMMH
10 ≤ WSMAX
WSMAX < 10
N = 71
$\bar{y}$ = 139.3
N = 6
$\bar{y}$ = 269.7
③
N = 46
$\bar{y}$ = 235.9
N = 3
$\bar{y}$ = 592.3
①
MXTMP < 80
80 ≤ MXTMP
PMMH < 1000
1000 ≤ PMMH
N: 43
$\bar{y}$ = 119.2
N = 28
$\bar{y}$ = 170.0
⑦
N = 11
$\bar{y}$ = 148.4
N = 35
$\bar{y}$ = 263.4
WV < 10
10 ≤ WV
.50 ≤ PREC3
.01 ≤ PREC3 < .50
MXTMP < 80
80 ≤ MXTMP
N: 14
$\bar{y}$ = 78.5
⑩
N = 29
$\bar{y}$ = 138.9
⑧
N = 12
$\bar{y}$ = 138.1
⑨
N = 16
$\bar{y}$ = 193.9
⑥
N = 11
$\bar{y}$ = 203.4
⑤
N = 24
$\bar{y}$ = 290.9
AVWS < 15
15 ≤ AVWS
N = 7
$\bar{y}$ = 231.6
④
N = 17
$\bar{y}$ = 315.5
②

Figure B6 - Decision tree for Stuttgart (#2), AR

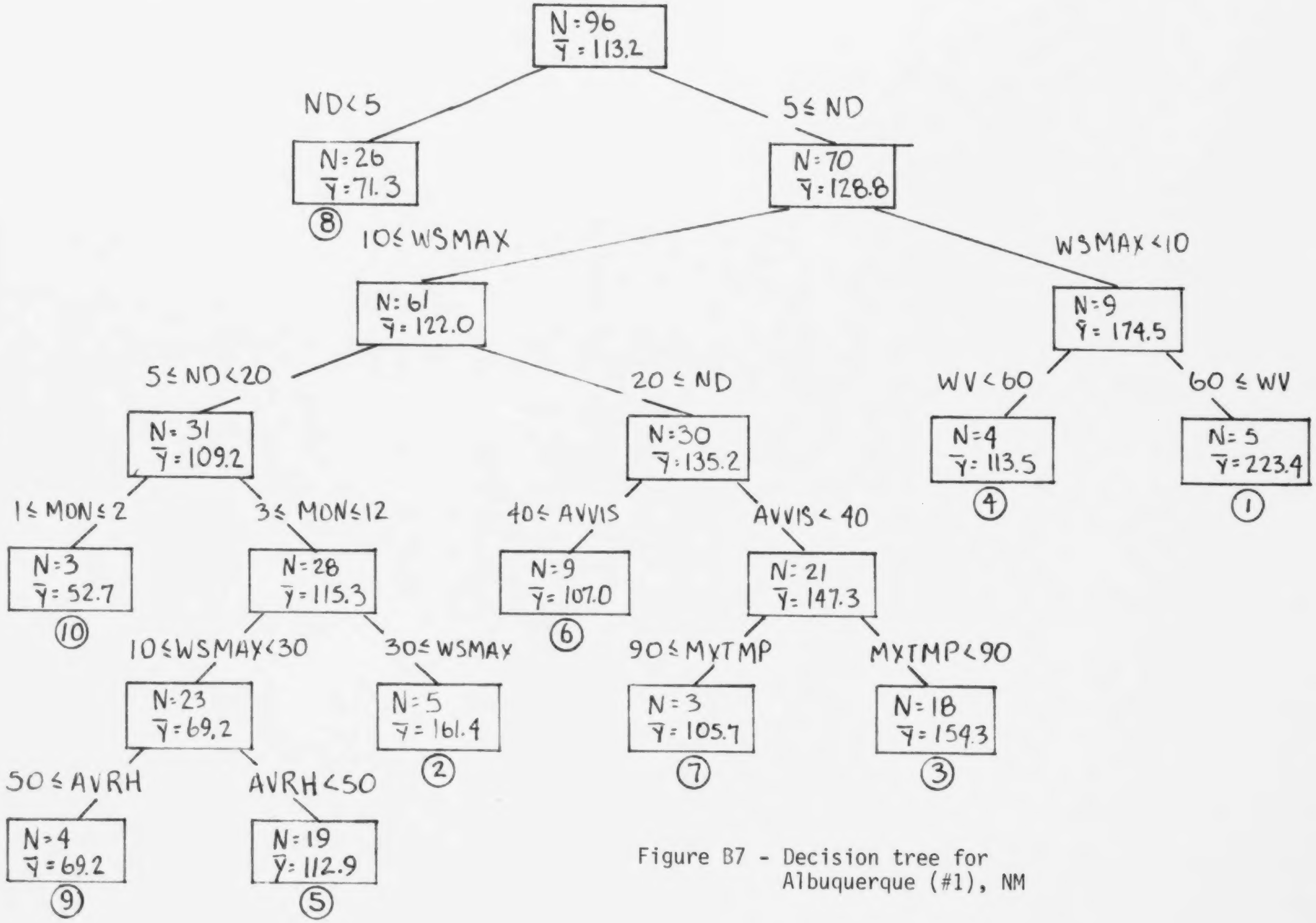

N:96
Ȳ = 113.2
ND<5
5≤ND
N:26
Ȳ:71.3
⑧
N:70
Ȳ:128.8
10≤WSMAX
WSMAX<10
N:61
Ȳ:122.0
N:9
Ȳ:174.5
5≤ND<20
20≤ND
WV<60
60≤WV
N:31
Ȳ:109.2
N:30
Ȳ:135.2
N:4
Ȳ:113.5
④
N:5
Ȳ:223.4
①
1≤MON≤2
3≤MON≤12
40≤AVVIS
AVVIS<40
N:3
Ȳ:52.7
⑩
N:28
Ȳ:115.3
N:9
Ȳ:107.0
⑥
N:21
Ȳ:147.3
10≤WSMAX<30
30≤WSMAX
90≤MXTMP
MXTMP<90
N:23
Ȳ:69.2
N:5
Ȳ:161.4
②
N:3
Ȳ:105.7
⑦
N:18
Ȳ:154.3
③
50≤AVRH
AVRH<50
N:4
Ȳ:69.2
⑨
N:19
Ȳ:112.9
⑤

Figure B7 - Decision tree for Albuquerque (#1), NM

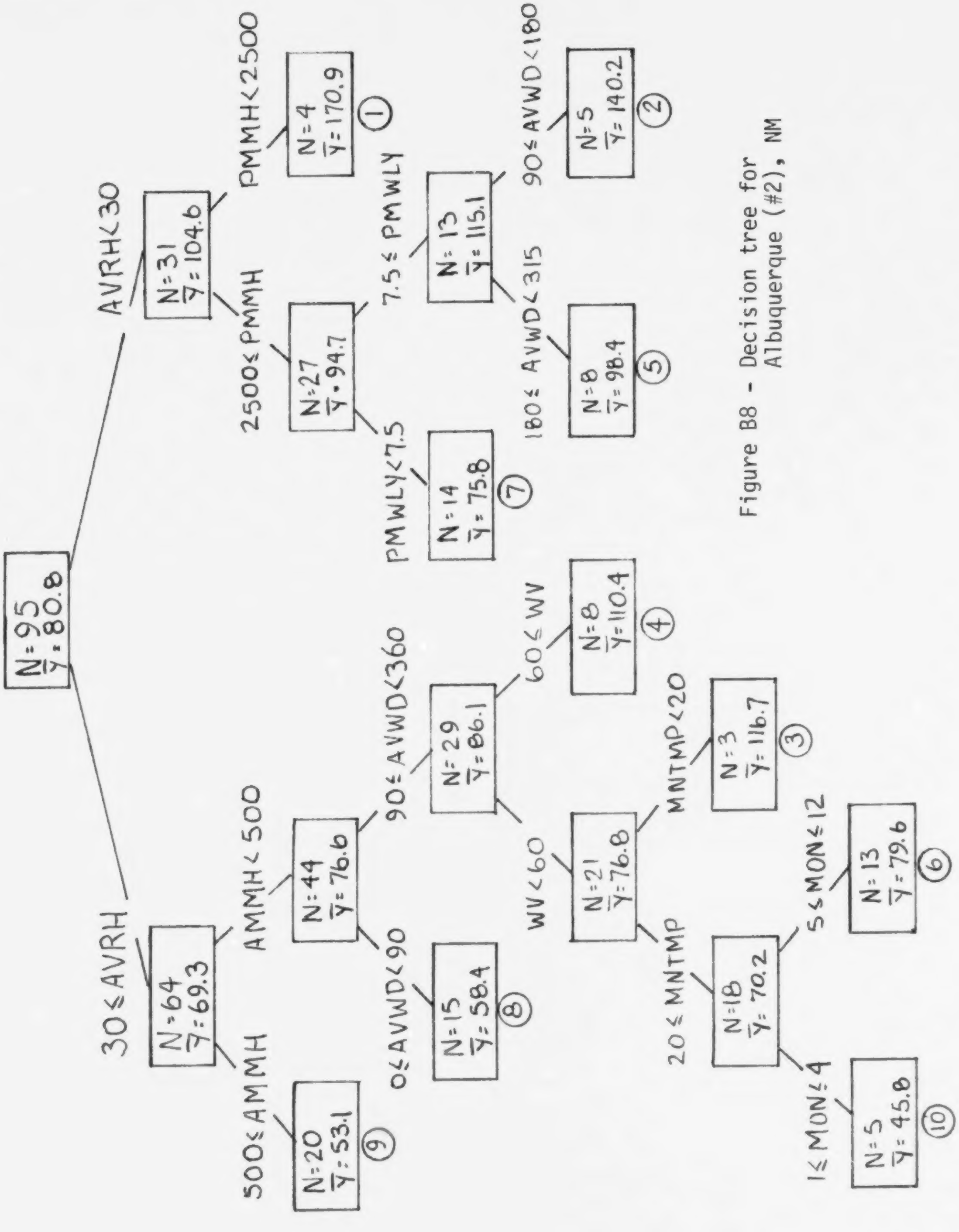
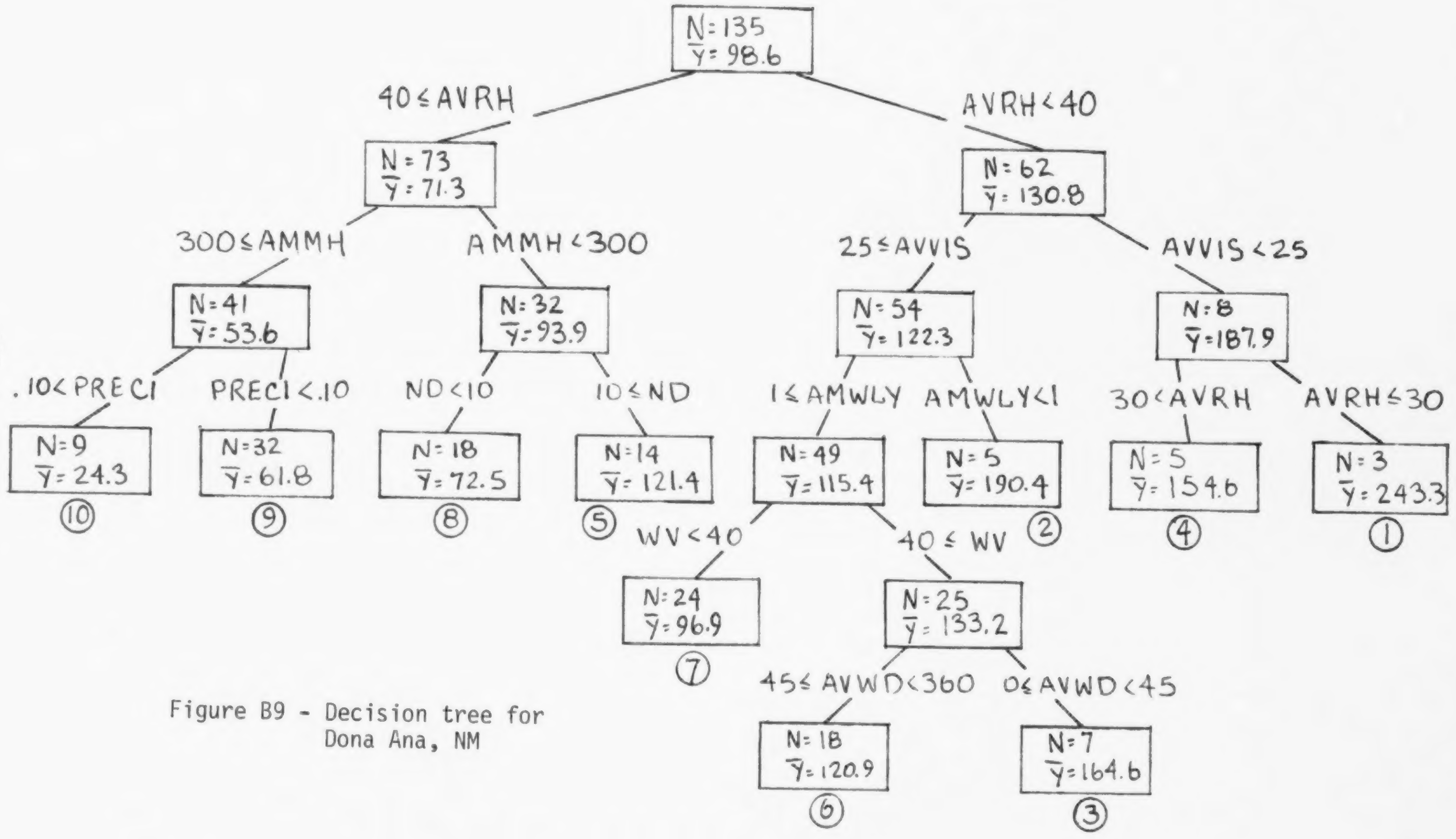

N=95
Ȳ=80.8
30≤AVRH
AVRH<30
N=64
Ȳ=69.3
N=31
Ȳ=104.6
500≤AMMH
AMMH<500
2500≤PMMH
PMMH<2500
N=20
Ȳ=53.1
⑨
N=44
Ȳ=76.6
N=27
Ȳ=94.7
N=4
Ȳ=170.9
①
0≤AVWD<90
90≤AVWD<360
PMWLY<7.5
7.5≤PMWLY
N=15
Ȳ=58.4
⑧
N=29
Ȳ=86.1
N=14
Ȳ=75.8
⑦
N=13
Ȳ=115.1
WV<60
60≤WV
180≤AVWD<315
90≤AVWD<180
N=21
Ȳ=76.8
N=8
Ȳ=110.4
④
N=8
Ȳ=98.4
⑤
N=5
Ȳ=140.2
②
20≤MNTMP
MNTMP<20
N=18
Ȳ=70.2
N=3
Ȳ=116.7
③
1≤MON≤4
5≤MON≤12
N=5
Ȳ=45.8
⑩
N=13
Ȳ=79.6
⑥

Figure B8 - Decision tree for Albuquerque (#2), NM



N=135
$\bar{y}$=98.6
40≤AVRH
AVRH<40
N=73
$\bar{y}$=71.3
N=62
$\bar{y}$=130.8
300≤AMMH
AMMH<300
25≤AVVIS
AVVIS<25
N=41
$\bar{y}$=53.6
N=32
$\bar{y}$=93.9
N=54
$\bar{y}$=122.3
N=8
$\bar{y}$=187.9
.10<PRECI
PRECI<.10
ND<10
10≤ND
1≤AMWLY
AMWLY<1
30<AVRH
AVRH≤30
N=9
$\bar{y}$=24.3
N=32
$\bar{y}$=61.8
N=18
$\bar{y}$=72.5
N=14
$\bar{y}$=121.4
N=49
$\bar{y}$=115.4
N=5
$\bar{y}$=190.4
N=5
$\bar{y}$=154.6
N=3
$\bar{y}$=243.3
⑩
⑨
⑧
⑤
②
④
①
WV<40
40≤WV
N=24
$\bar{y}$=96.9
N=25
$\bar{y}$=133.2
⑦
45≤AVWD<360
0≤AVWD<45
N=18
$\bar{y}$=120.9
N=7
$\bar{y}$=164.6
⑥
③

Figure B9 - Decision tree for
Dona Ana, NM

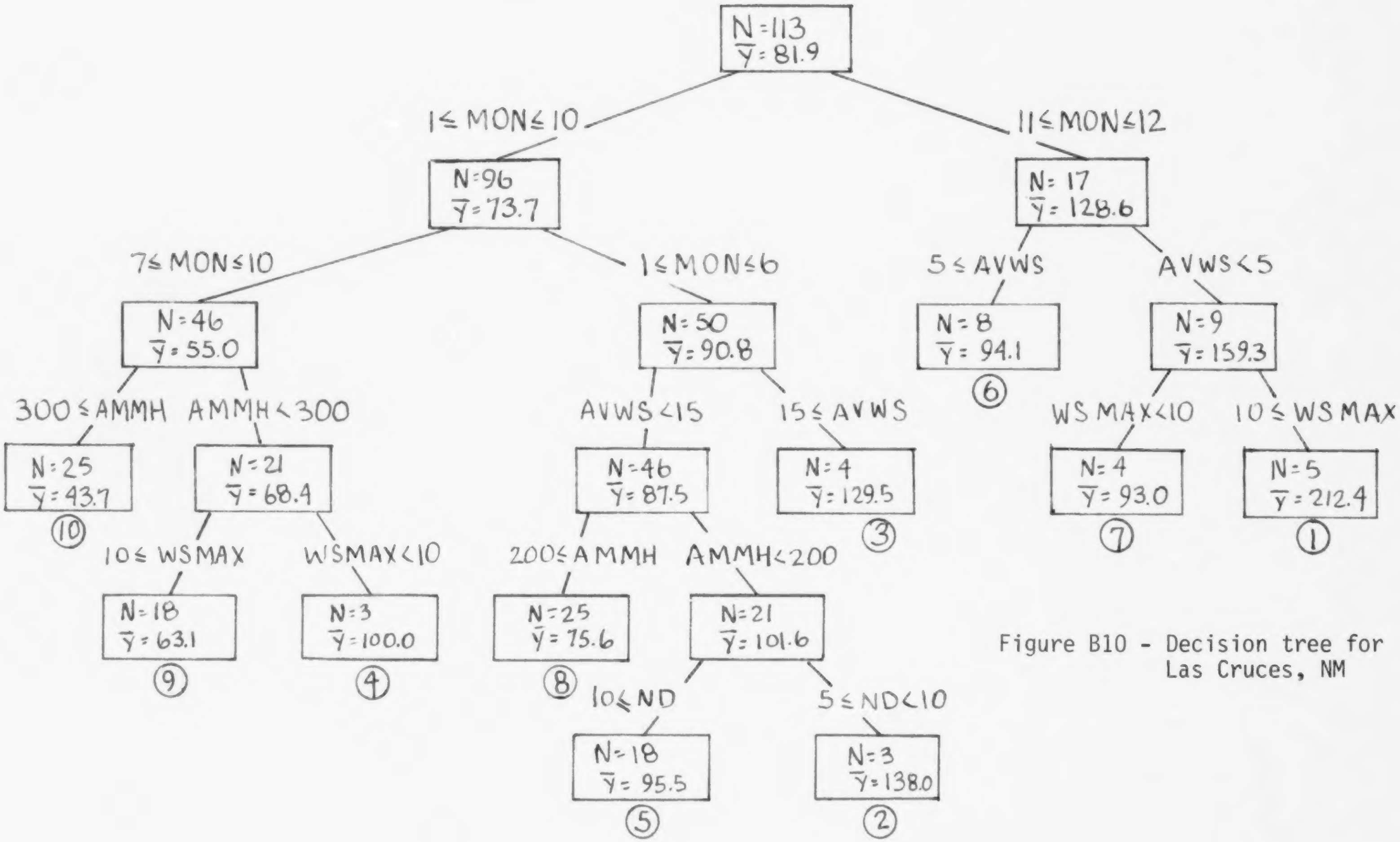

N=113
$\bar{y}$=81.9
1≤MON≤10
11≤MON≤12
N=96
$\bar{y}$=73.7
N=17
$\bar{y}$=128.6
7≤MON≤10
1≤MON≤6
5≤AVWS
AVWS<5
N=46
$\bar{y}$=55.0
N=50
$\bar{y}$=90.8
N=8
$\bar{y}$=94.1
⑥
N=9
$\bar{y}$=159.3
300≤AMMH
AMMH<300
AVWS<15
15≤AVWS
WSMAX<10
10≤WSMAX
N=25
$\bar{y}$=43.7
⑩
N=21
$\bar{y}$=68.4
N=46
$\bar{y}$=87.5
N=4
$\bar{y}$=129.5
③
N=4
$\bar{y}$=93.0
⑦
N=5
$\bar{y}$=212.4
①
10≤WSMAX
WSMAX<10
200≤AMMH
AMMH<200
N=18
$\bar{y}$=63.1
⑨
N=3
$\bar{y}$=100.0
④
N=25
$\bar{y}$=75.6
⑧
N=21
$\bar{y}$=101.6
10≤ND
5≤ND<10
N=18
$\bar{y}$=95.5
⑤
N=3
$\bar{y}$=138.0
②

Figure B10 - Decision tree for Las Cruces, NM

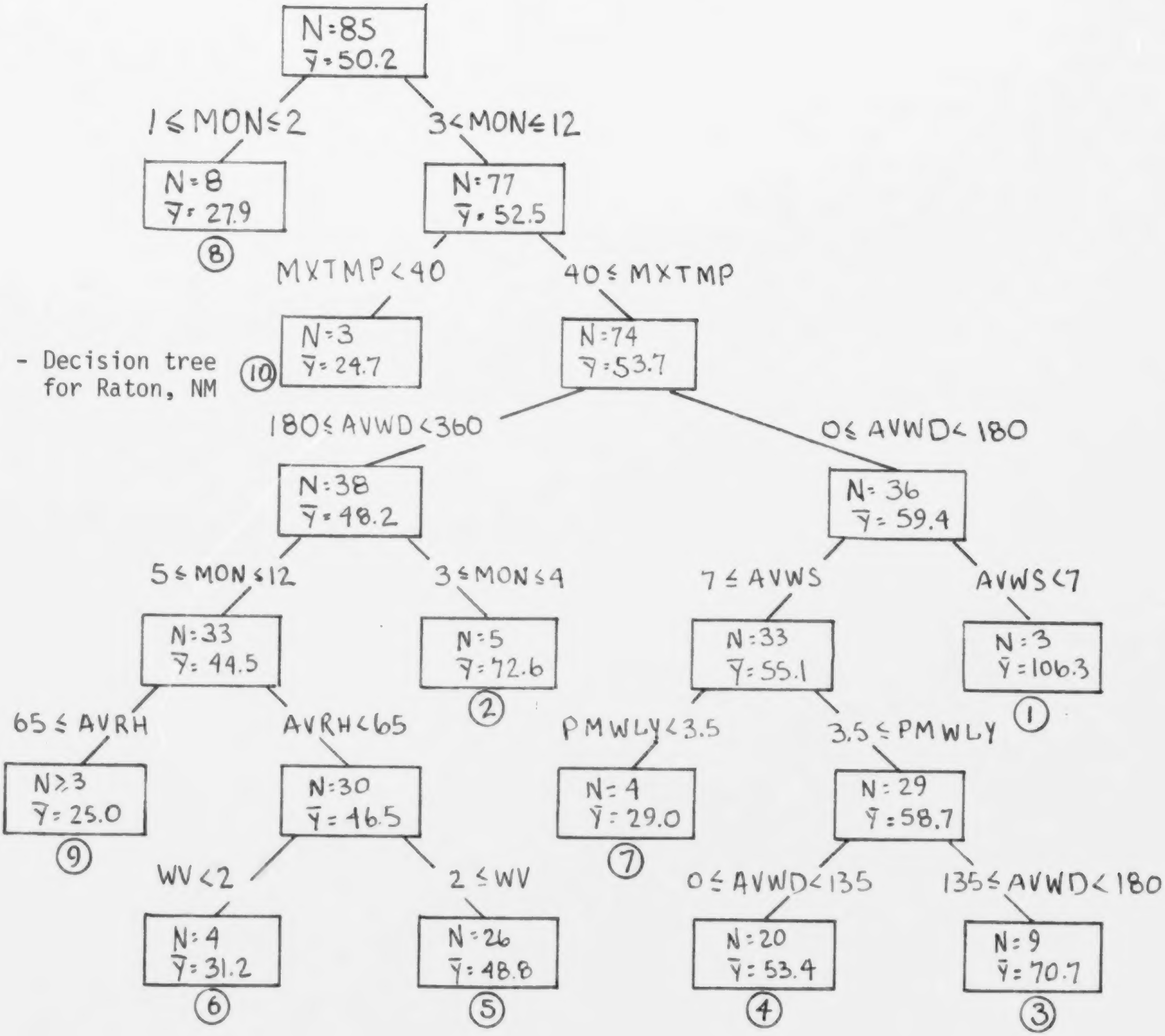

N=85
$\bar{Y}$=50.2
1≤MON≤2
3<MON≤12
N=8
$\bar{Y}$=27.9
⑧
N=77
$\bar{Y}$=52.5
MXTMP<40
40≤MXTMP
N=3
$\bar{Y}$=24.7
⑩
N=74
$\bar{Y}$=53.7
180≤AVWD<360
0≤AVWD<180
N=38
$\bar{Y}$=48.2
N=36
$\bar{Y}$=59.4
5≤MON≤12
3≤MON≤4
7≤AVWS
AVWS<7
N=33
$\bar{Y}$=44.5
N=5
$\bar{Y}$=72.6
②
N=33
$\bar{Y}$=55.1
N=3
$\bar{Y}$=106.3
①
65≤AVRH
AVRH<65
PMWLY<3.5
3.5≤PMWLY
N≥3
$\bar{Y}$=25.0
⑨
N=30
$\bar{Y}$=46.5
N=4
$\bar{Y}$=29.0
⑦
N=29
$\bar{Y}$=58.7
WV<2
2≤WV
0≤AVWD<135
135≤AVWD<180
N=4
$\bar{Y}$=31.2
⑥
N=26
$\bar{Y}$=48.8
⑤
N=20
$\bar{Y}$=53.4
④
N=9
$\bar{Y}$=70.7
③

Figure B11 - Decision tree for Raton, NM

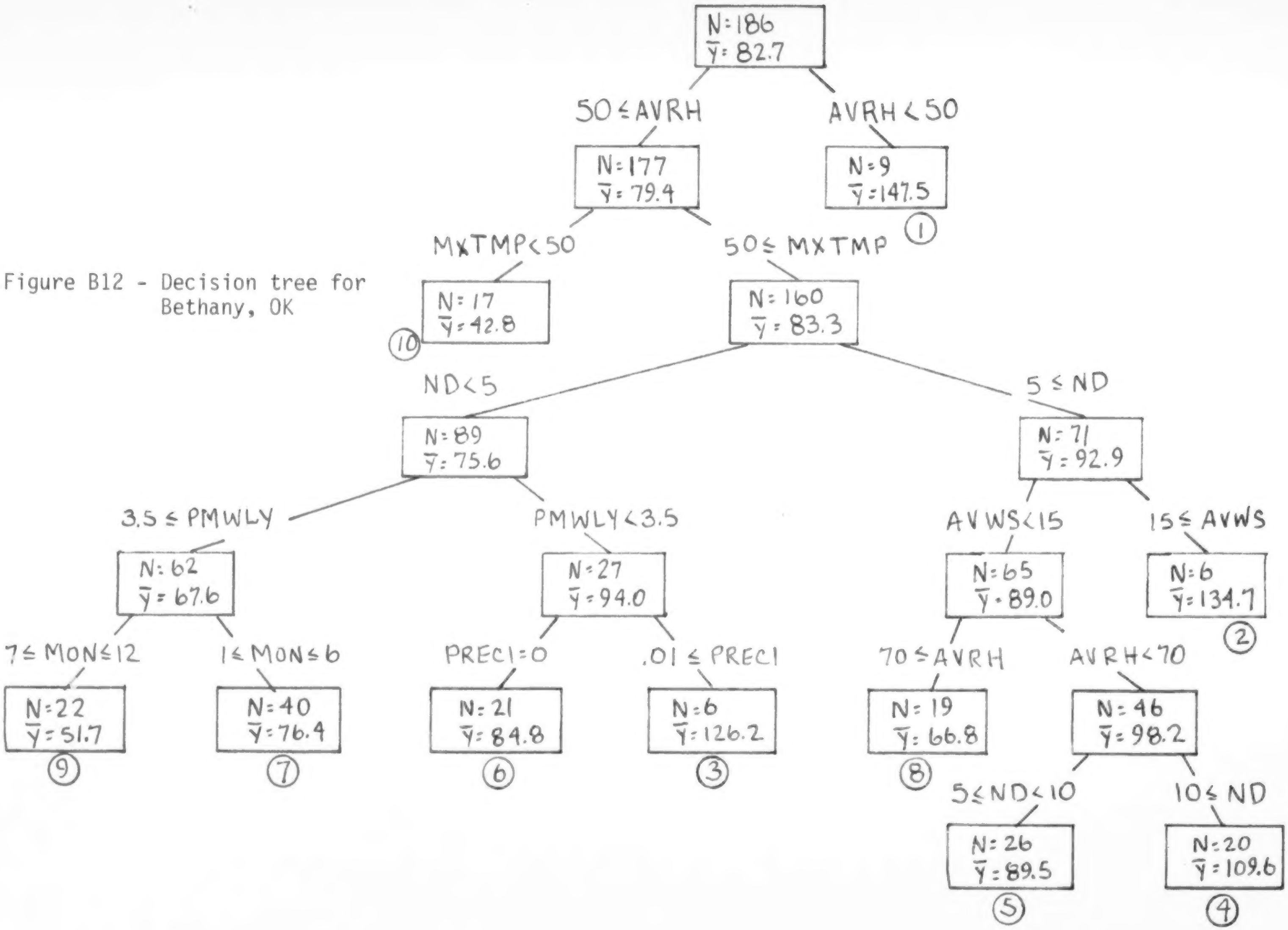

N=186
$\bar{y}$=82.7
50≤AVRH
AVRH<50
N=177
$\bar{y}$=79.4
N=9
$\bar{y}$=147.5
①
MXTMP<50
50≤MXTMP
N=17
$\bar{y}$=42.8
⑩
N=160
$\bar{y}$=83.3
ND<5
5≤ND
N=89
$\bar{y}$=75.6
N=71
$\bar{y}$=92.9
3.5≤PMWLY
PMWLY<3.5
AVWS<15
15≤AVWS
N=62
$\bar{y}$=67.6
N=27
$\bar{y}$=94.0
N=65
$\bar{y}$=89.0
N=6
$\bar{y}$=134.7
②
7≤MON≤12
1≤MON≤6
PRECI=0
.01≤PRECI
70≤AVRH
AVRH<70
N=22
$\bar{y}$=51.7
⑨
N=40
$\bar{y}$=76.4
⑦
N=21
$\bar{y}$=84.8
⑥
N=6
$\bar{y}$=126.2
③
N=19
$\bar{y}$=66.8
⑧
N=46
$\bar{y}$=98.2
5≤ND<10
10≤ND
N=26
$\bar{y}$=89.5
⑤
N=20
$\bar{y}$=109.6
④

Figure B12 - Decision tree for
Bethany, OK

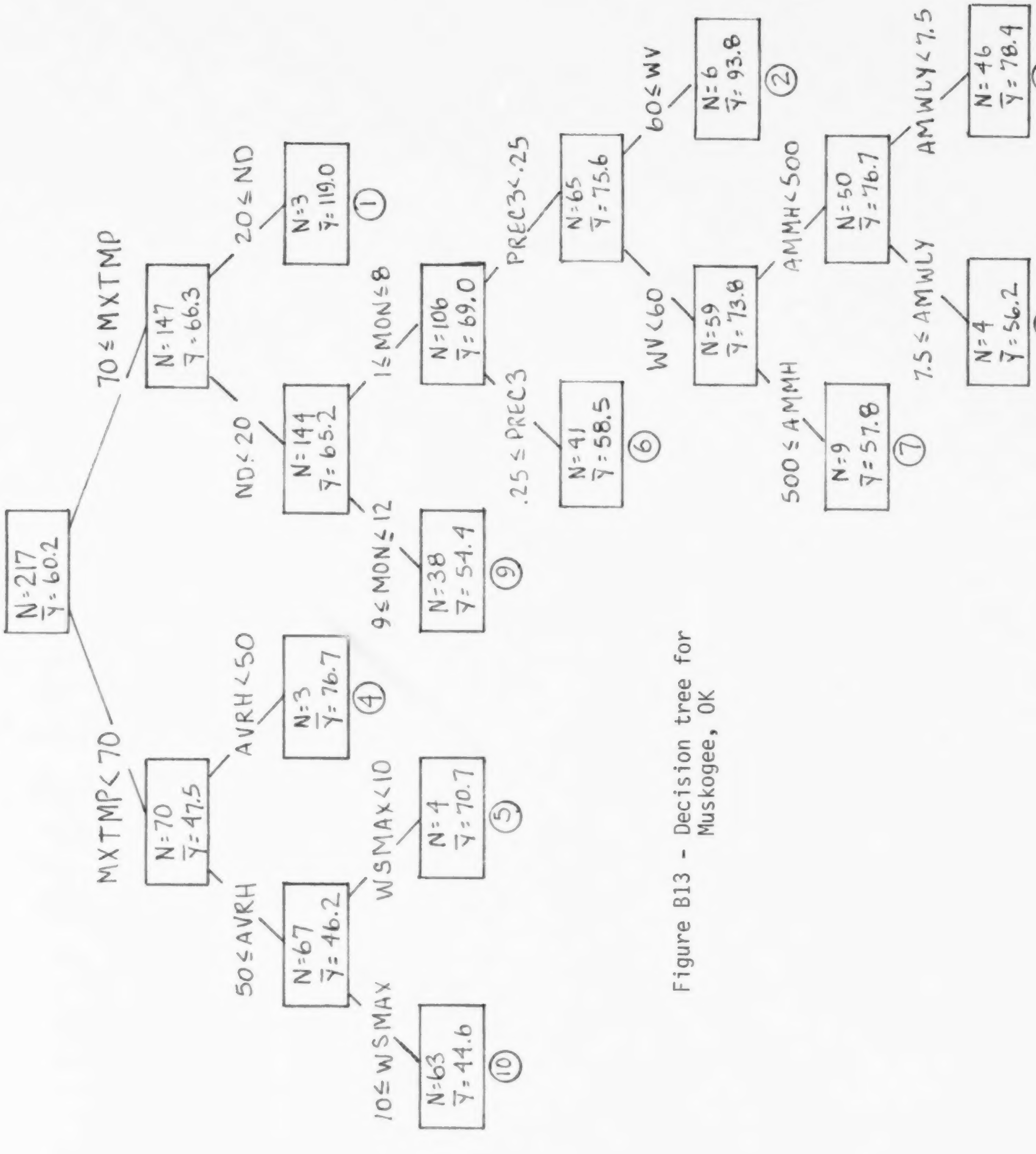

N=217
Ȳ=60.2
MXTMP<70
70≤MXTMP
N=70
Ȳ=47.5
N=147
Ȳ=66.3
50≤AVRH
AVRH<50
ND<20
20≤ND
N=67
Ȳ=46.2
N=3
Ȳ=76.7
④
N=144
Ȳ=65.2
N=3
Ȳ=119.0
①
10≤WSMAX
WSMAX<10
9≤MON≤12
1≤MON≤8
N=63
Ȳ=44.6
⑩
N=4
Ȳ=70.7
⑤
N=38
Ȳ=54.4
⑨
N=106
Ȳ=69.0
.25≤PREC3
PREC3<.25
N=41
Ȳ=58.5
⑥
N=65
Ȳ=75.6
WV<60
60≤WV
N=59
Ȳ=73.8
N=6
Ȳ=93.8
②
500≤AMMH
AMMH<500
N=9
Ȳ=57.8
⑦
N=50
Ȳ=76.7
7.5≤AMWLY
AMWLY<7.5
N=4
Ȳ=56.2
⑧
N=46
Ȳ=78.4
③

Figure B13 - Decision tree for Muskogee, OK

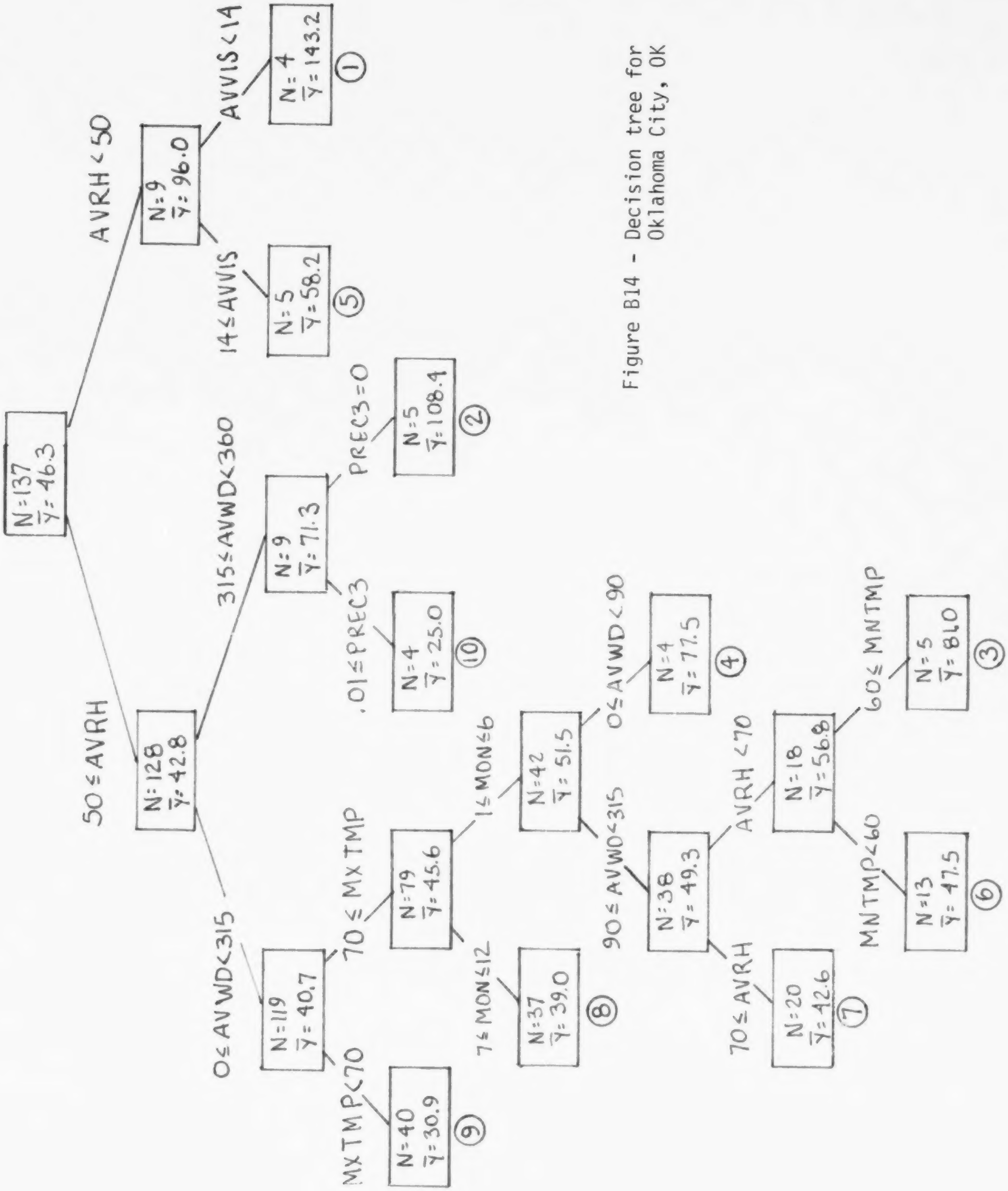

N=137
Ȳ=46.3
AVRH<50
N=9
Ȳ=96.0
AVVIS<14
N=4
Ȳ=143.2
①
14≤AVVIS
N=5
Ȳ=58.2
⑤
50≤AVRH
N=128
Ȳ=42.8
315≤AVWD<360
N=9
Ȳ=71.3
PREC3=0
N=5
Ȳ=108.4
②
.01≤PREC3
N=4
Ȳ=25.0
⑩
0≤AVWD<315
N=119
Ȳ=40.7
70≤MXTMP
N=79
Ȳ=45.6
1≤MON≤6
N=42
Ȳ=51.5
0≤AVWD<90
N=4
Ȳ=77.5
④
90≤AVWD<315
N=38
Ȳ=49.3
AVRH<70
N=18
Ȳ=56.8
60≤MNTMP
N=5
Ȳ=81.0
③
MNTMP<60
N=13
Ȳ=47.5
⑥
70≤AVRH
N=20
Ȳ=42.6
⑦
7≤MON≤12
N=37
Ȳ=39.0
⑧
MXTMP<70
N=40
Ȳ=30.9
⑨

Figure B14 - Decision tree for Oklahoma City, OK

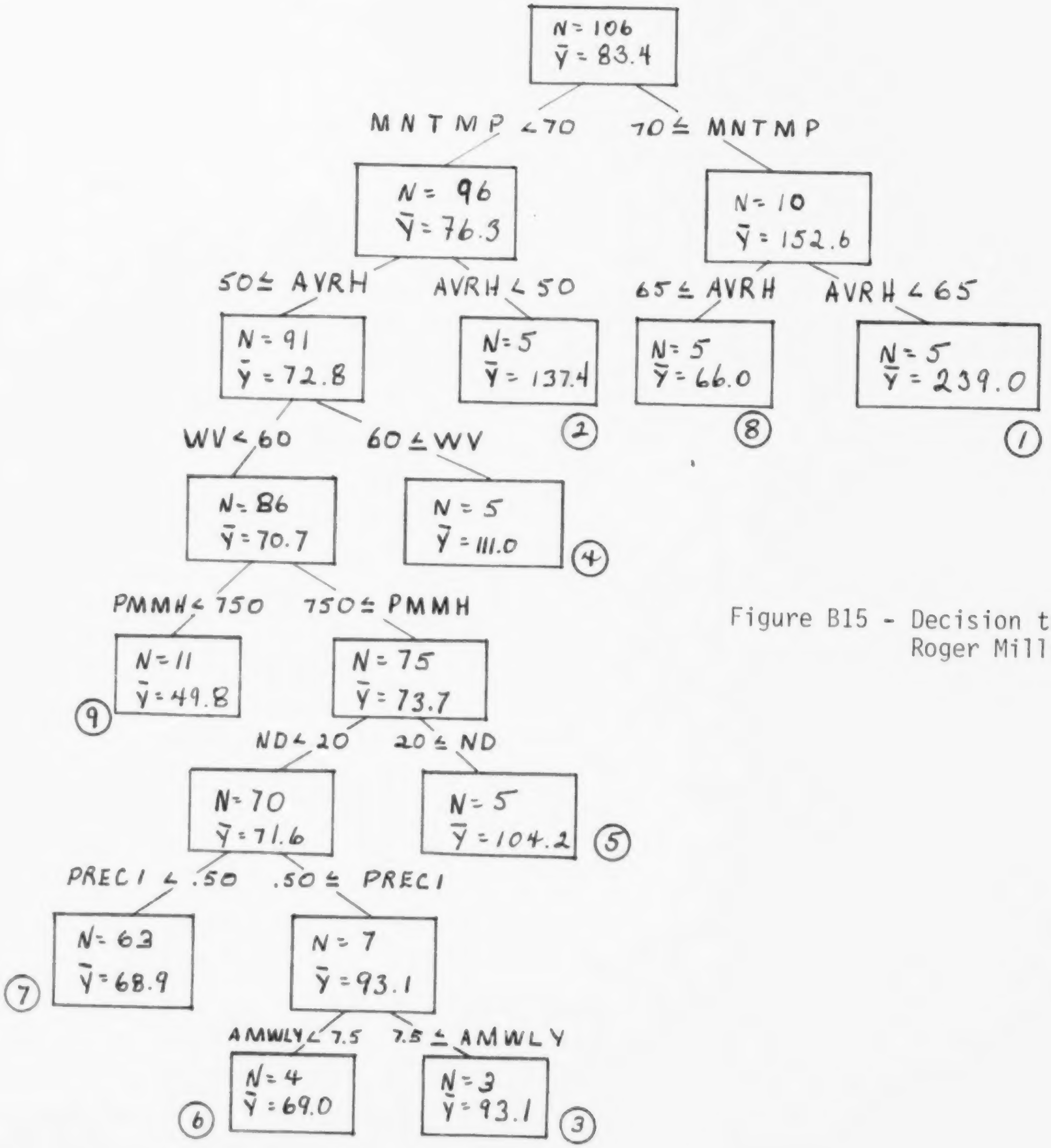

N = 106
Ȳ = 83.4
MNTMP < 70
70 ≤ MNTMP
N = 96
Ȳ = 76.3
N = 10
Ȳ = 152.6
50 ≤ AVRH
AVRH < 50
65 ≤ AVRH
AVRH < 65
N = 91
Ȳ = 72.8
N = 5
Ȳ = 137.4
②
N = 5
Ȳ = 66.0
⑧
N = 5
Ȳ = 239.0
①
WV < 60
60 ≤ WV
N = 86
Ȳ = 70.7
N = 5
Ȳ = 111.0
④
PMMH < 750
750 ≤ PMMH
N = 11
Ȳ = 49.8
⑨
N = 75
Ȳ = 73.7
ND < 20
20 ≤ ND
N = 70
Ȳ = 71.6
N = 5
Ȳ = 104.2
⑤
PRECI < .50
.50 ≤ PRECI
N = 63
Ȳ = 68.9
⑦
N = 7
Ȳ = 93.1
AMWLY < 7.5
7.5 ≤ AMWLY
N = 4
Ȳ = 69.0
⑥
N = 3
Ȳ = 93.1
③

Figure B15 - Decision tree for
Roger Mills, OK

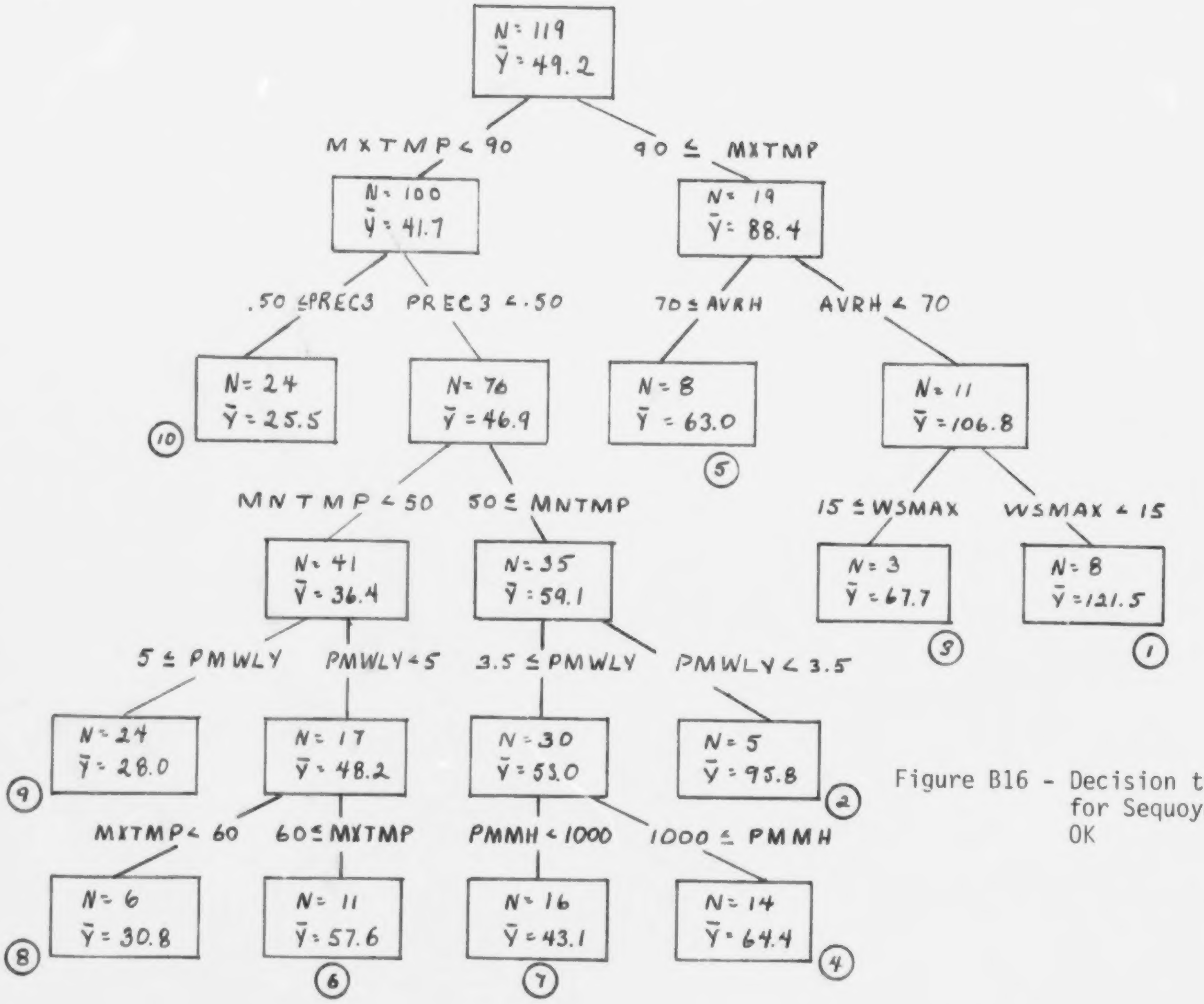

N= 119
Ȳ = 49.2
MXTMP < 90
90 ≤ MXTMP
N= 100
Ȳ = 41.7
N= 19
Ȳ = 88.4
.50 ≤ PREC3
PREC3 < .50
70 ≤ AVRH
AVRH < 70
N= 24
Ȳ = 25.5
⑩
N= 76
Ȳ = 46.9
N= 8
Ȳ = 63.0
⑤
N= 11
Ȳ = 106.8
MNTMP < 50
50 ≤ MNTMP
15 ≤ WSMAX
WSMAX < 15
N= 41
Ȳ = 36.4
N= 35
Ȳ = 59.1
N= 3
Ȳ = 67.7
③
N= 8
Ȳ = 121.5
①
5 ≤ PMWLY
PMWLY < 5
3.5 ≤ PMWLY
PMWLY < 3.5
N= 24
Ȳ = 28.0
⑨
N= 17
Ȳ = 48.2
N= 30
Ȳ = 53.0
N= 5
Ȳ = 95.8
②
MXTMP < 60
60 ≤ MXTMP
PMMH < 1000
1000 ≤ PMMH
N= 6
Ȳ = 30.8
⑧
N= 11
Ȳ = 57.6
⑥
N= 16
Ȳ = 43.1
⑦
N= 14
Ȳ = 64.4
④

Figure B16 - Decision tree for Sequoyah Co., OK

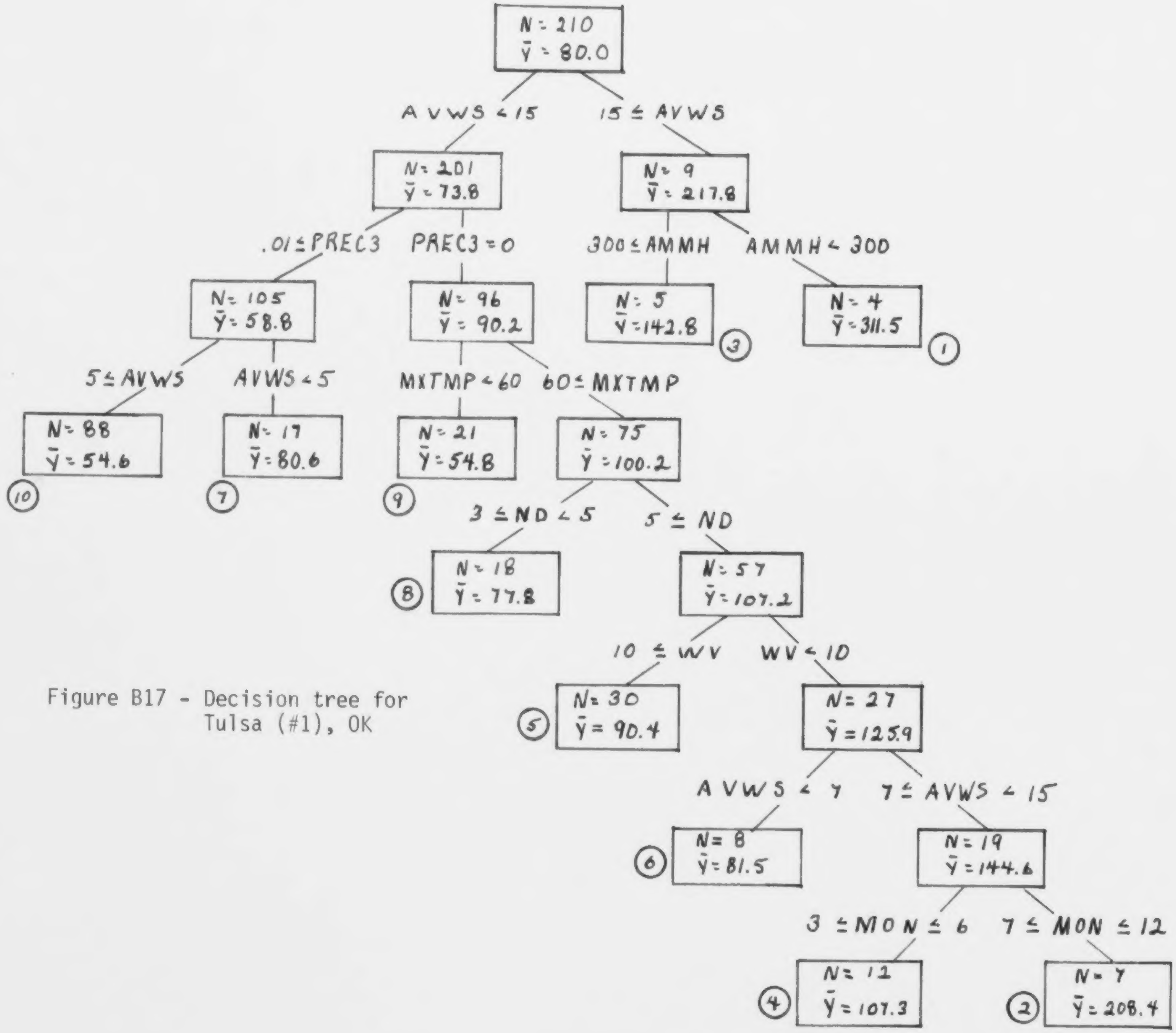

N = 210
Ȳ = 80.0
AVWS < 15
15 ≤ AVWS
N = 201
Ȳ = 73.8
N = 9
Ȳ = 217.8
.01 ≤ PREC3
PREC3 = 0
300 ≤ AMMH
AMMH < 300
N = 105
Ȳ = 58.8
N = 96
Ȳ = 90.2
N = 5
Ȳ = 142.8
③
N = 4
Ȳ = 311.5
①
5 ≤ AVWS
AVWS < 5
MXTMP < 60
60 ≤ MXTMP
N = 88
Ȳ = 54.6
⑩
N = 17
Ȳ = 80.6
⑦
N = 21
Ȳ = 54.8
⑨
N = 75
Ȳ = 100.2
3 ≤ ND < 5
5 ≤ ND
N = 18
Ȳ = 77.8
⑧
N = 57
Ȳ = 107.2
10 ≤ WV
WV < 10
N = 30
Ȳ = 90.4
⑤
N = 27
Ȳ = 125.9
AVWS < 7
7 ≤ AVWS < 15
N = 8
Ȳ = 81.5
⑥
N = 19
Ȳ = 144.6
3 ≤ MON ≤ 6
7 ≤ MON ≤ 12
N = 12
Ȳ = 107.3
④
N = 7
Ȳ = 208.4
②

Figure B17 - Decision tree for Tulsa (#1), OK

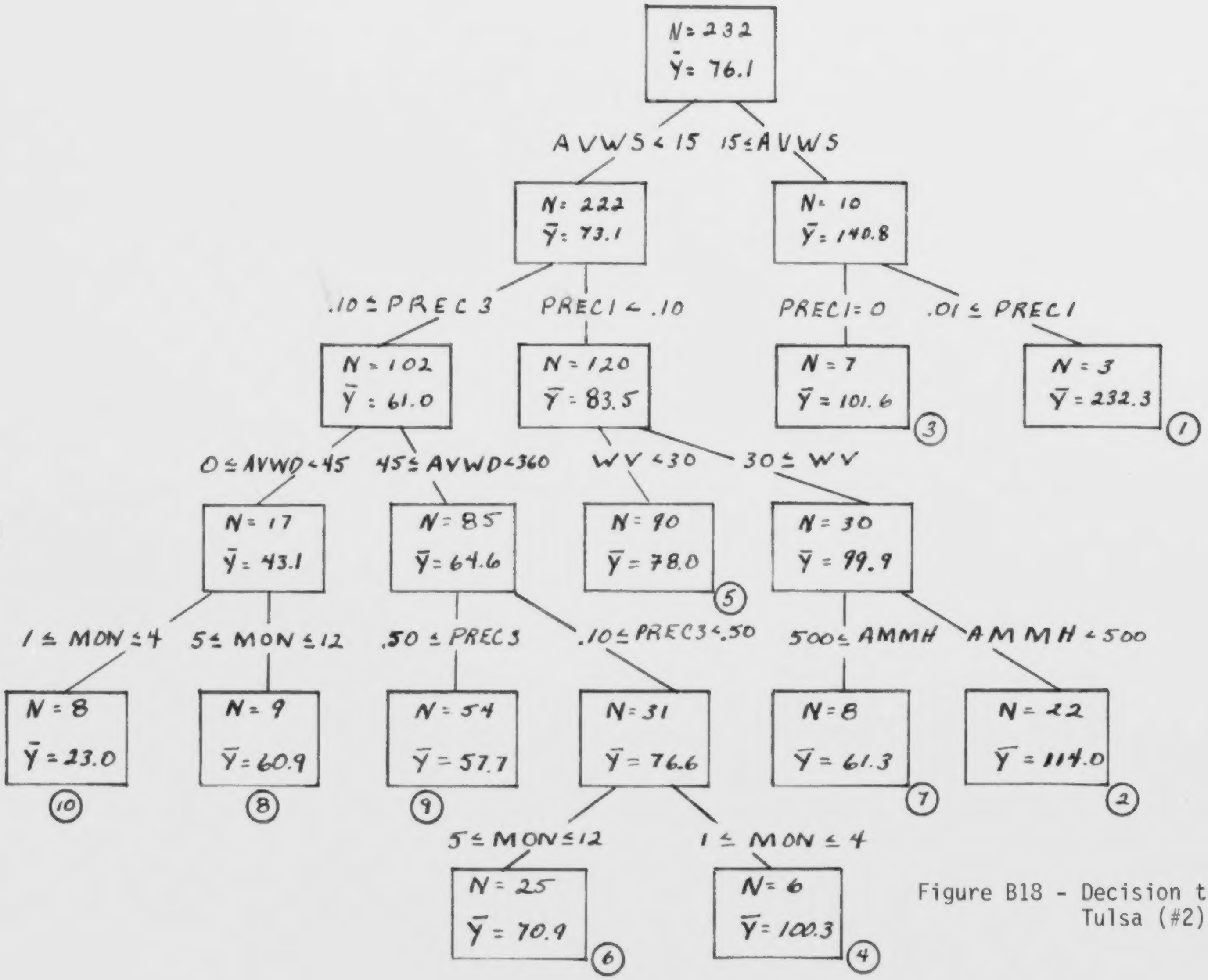

N=232
Ȳ= 76.1
AVWS < 15
15 ≤ AVWS
N= 222
Ȳ= 73.1
N= 10
Ȳ= 140.8
.10 ≤ PREC 3
PREC1 < .10
PREC1 = 0
.01 ≤ PREC1
N = 102
Ȳ = 61.0
N = 120
Ȳ = 83.5
N = 7
Ȳ = 101.6
③
N = 3
Ȳ = 232.3
①
0 ≤ AVWD < 45
45 ≤ AVWD < 360
WV < 30
30 ≤ WV
N= 17
Ȳ= 43.1
N= 85
Ȳ= 64.6
N= 90
Ȳ = 78.0
⑤
N= 30
Ȳ = 99.9
1 ≤ MON ≤ 4
5 ≤ MON ≤ 12
.50 ≤ PREC3
.10 ≤ PREC3 < .50
500 ≤ AMMH
AMMH < 500
N = 8
Ȳ = 23.0
⑩
N = 9
Ȳ = 60.9
⑧
N = 54
Ȳ = 57.7
⑨
N = 31
Ȳ = 76.6
N = 8
Ȳ = 61.3
⑦
N = 22
Ȳ = 114.0
②
5 ≤ MON ≤ 12
1 ≤ MON ≤ 4
N = 25
Ȳ = 70.9
⑥
N = 6
Ȳ = 100.3
④

Figure B18 - Decision tree for Tulsa (#2), OK

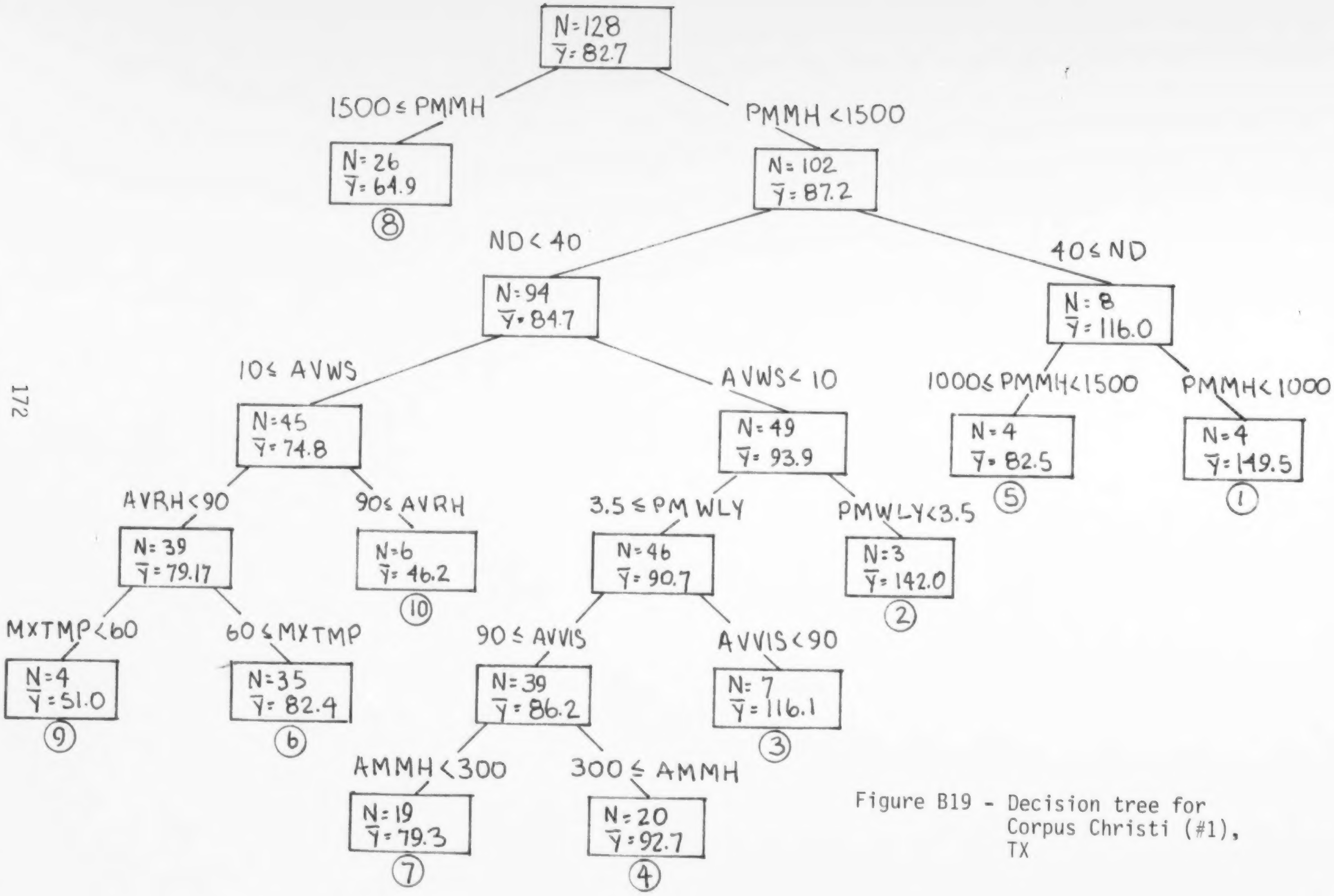

N=128
Ȳ=82.7
1500≤PMMH
PMMH<1500
N=26
Ȳ=64.9
⑧
N=102
Ȳ=87.2
ND<40
40≤ND
N=94
Ȳ=84.7
N=8
Ȳ=116.0
10≤AVWS
AVWS<10
1000≤PMMH<1500
PMMH<1000
N=45
Ȳ=74.8
N=49
Ȳ=93.9
N=4
Ȳ=82.5
⑤
N=4
Ȳ=149.5
①
AVRH<90
90≤AVRH
3.5≤PMWLY
PMWLY<3.5
N=39
Ȳ=79.17
N=6
Ȳ=46.2
⑩
N=46
Ȳ=90.7
N=3
Ȳ=142.0
②
MXTMP<60
60≤MXTMP
90≤AVVIS
AVVIS<90
N=4
Ȳ=51.0
⑨
N=35
Ȳ=82.4
⑥
N=39
Ȳ=86.2
N=7
Ȳ=116.1
③
AMMH<300
300≤AMMH
N=19
Ȳ=79.3
⑦
N=20
Ȳ=92.7
④

Figure B19 - Decision tree for Corpus Christi (#1), TX

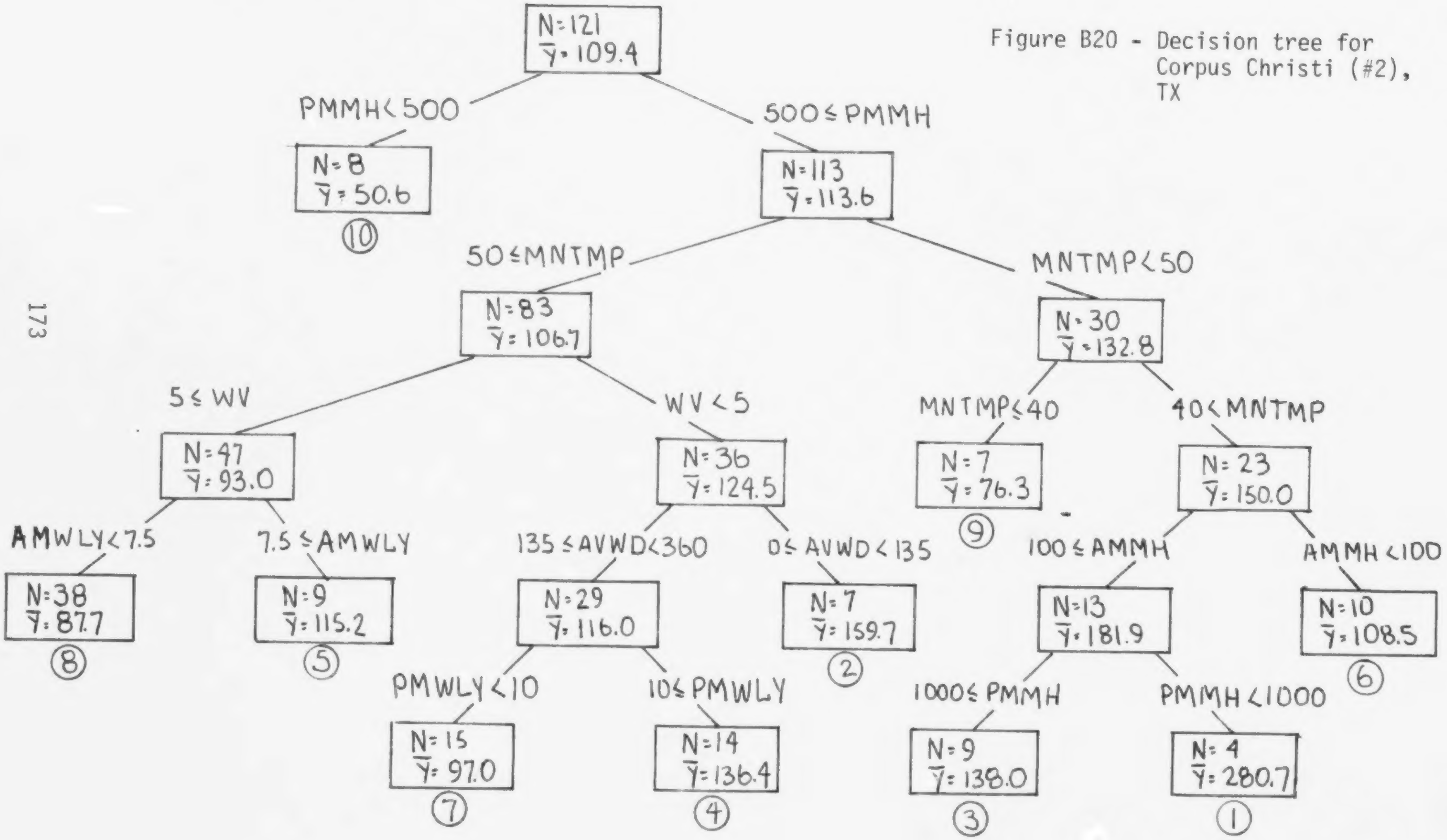

N=121
$\bar{y}$=109.4
PMMH<500
500≤PMMH
N=8
$\bar{y}$=50.6
⑩
N=113
$\bar{y}$=113.6
50≤MNTMP
MNTMP<50
N=83
$\bar{y}$=106.7
N=30
$\bar{y}$=132.8
5≤WV
WV<5
MNTMP≤40
40<MNTMP
N=47
$\bar{y}$=93.0
N=36
$\bar{y}$=124.5
N=7
$\bar{y}$=76.3
⑨
N=23
$\bar{y}$=150.0
AMWLY<7.5
7.5≤AMWLY
135≤AVWD<360
0≤AVWD<135
100≤AMMH
AMMH<100
N=38
$\bar{y}$=87.7
⑧
N=9
$\bar{y}$=115.2
⑤
N=29
$\bar{y}$=116.0
N=7
$\bar{y}$=159.7
②
N=13
$\bar{y}$=181.9
N=10
$\bar{y}$=108.5
⑥
PMWLY<10
10≤PMWLY
1000≤PMMH
PMMH<1000
N=15
$\bar{y}$=97.0
⑦
N=14
$\bar{y}$=136.4
④
N=9
$\bar{y}$=138.0
③
N=4
$\bar{y}$=280.7
①

Figure B20 - Decision tree for Corpus Christi (#2), TX

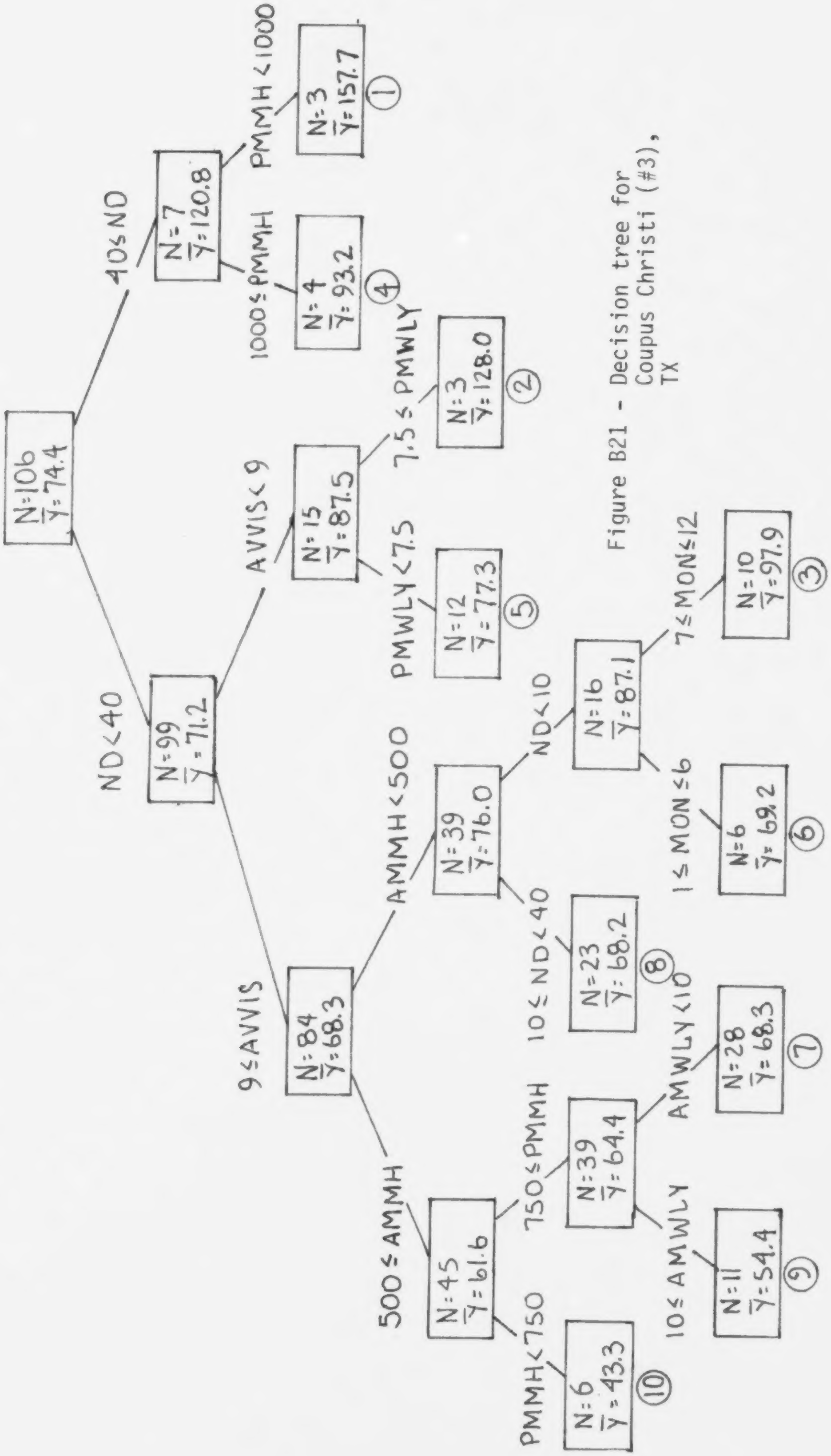

N:106
$\bar{Y}$: 74.4
ND<40
40≤ND
N:99
$\bar{Y}$: 71.2
N:7
$\bar{Y}$: 120.8
1000≤PMMH
PMMH<1000
N:4
$\bar{Y}$: 93.2
④
N:3
$\bar{Y}$: 157.7
①
9≤AVVIS
AVVIS<9
N:15
$\bar{Y}$: 87.5
PMWLY<7.5
7.5≤PMWLY
N:12
$\bar{Y}$: 77.3
⑤
N:3
$\bar{Y}$: 128.0
②
N:84
$\bar{Y}$: 68.3
500≤AMMH
AMMH<500
N:39
$\bar{Y}$: 76.0
10≤ND<40
ND<10
N:23
$\bar{Y}$: 68.2
⑧
N:16
$\bar{Y}$: 87.1
1≤MON≤6
7≤MON≤12
N:6
$\bar{Y}$: 69.2
⑥
N:10
$\bar{Y}$: 97.9
③
N:45
$\bar{Y}$: 61.6
PMMH<750
750≤PMMH
N:6
$\bar{Y}$: 43.3
⑩
N:39
$\bar{Y}$: 64.4
10≤AMWLY
AMWLY<10
N:11
$\bar{Y}$: 54.4
⑨
N:28
$\bar{Y}$: 68.3
⑦

Figure B21 - Decision tree for Coupus Christi (#3), TX

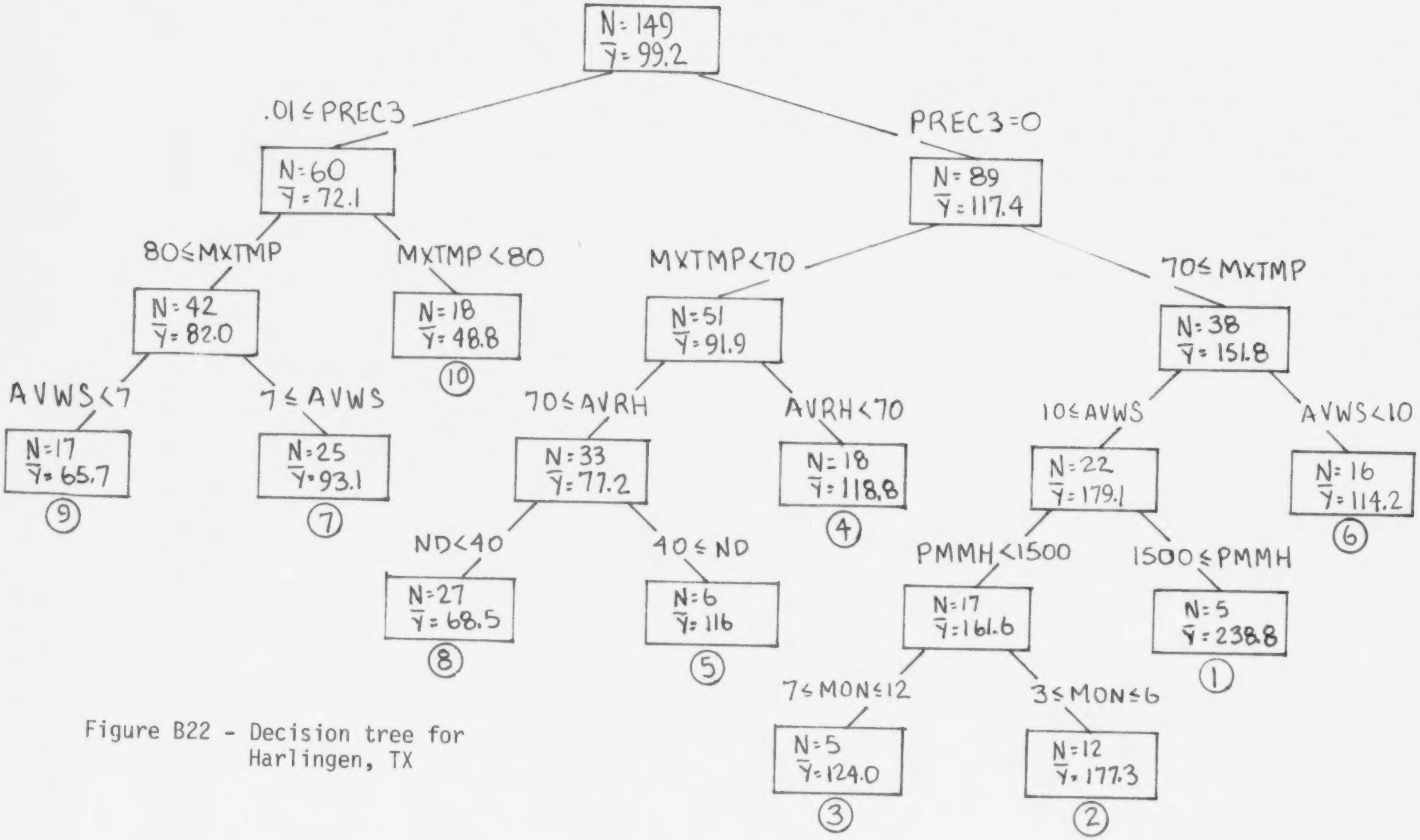

N: 149
Ȳ: 99.2
.01 ≤ PREC3
PREC3 = 0
N: 60
Ȳ: 72.1
N: 89
Ȳ: 117.4
80 ≤ MXTMP
MXTMP < 80
MXTMP < 70
70 ≤ MXTMP
N: 42
Ȳ: 82.0
N: 18
Ȳ: 48.8
⑩
N: 51
Ȳ: 91.9
N: 38
Ȳ: 151.8
AVWS < 7
7 ≤ AVWS
70 ≤ AVRH
AVRH < 70
10 ≤ AVWS
AVWS < 10
N: 17
Ȳ: 65.7
⑨
N: 25
Ȳ: 93.1
⑦
N: 33
Ȳ: 77.2
N: 18
Ȳ: 118.8
④
N: 22
Ȳ: 179.1
N: 16
Ȳ: 114.2
⑥
ND < 40
40 ≤ ND
PMMH < 1500
1500 ≤ PMMH
N: 27
Ȳ: 68.5
⑧
N: 6
Ȳ: 116
⑤
N: 17
Ȳ: 161.6
N: 5
Ȳ: 238.8
①
7 ≤ MON ≤ 12
3 ≤ MON ≤ 6
N: 5
Ȳ: 124.0
③
N: 12
Ȳ: 177.3
②

Figure B22 - Decision tree for Harlingen, TX

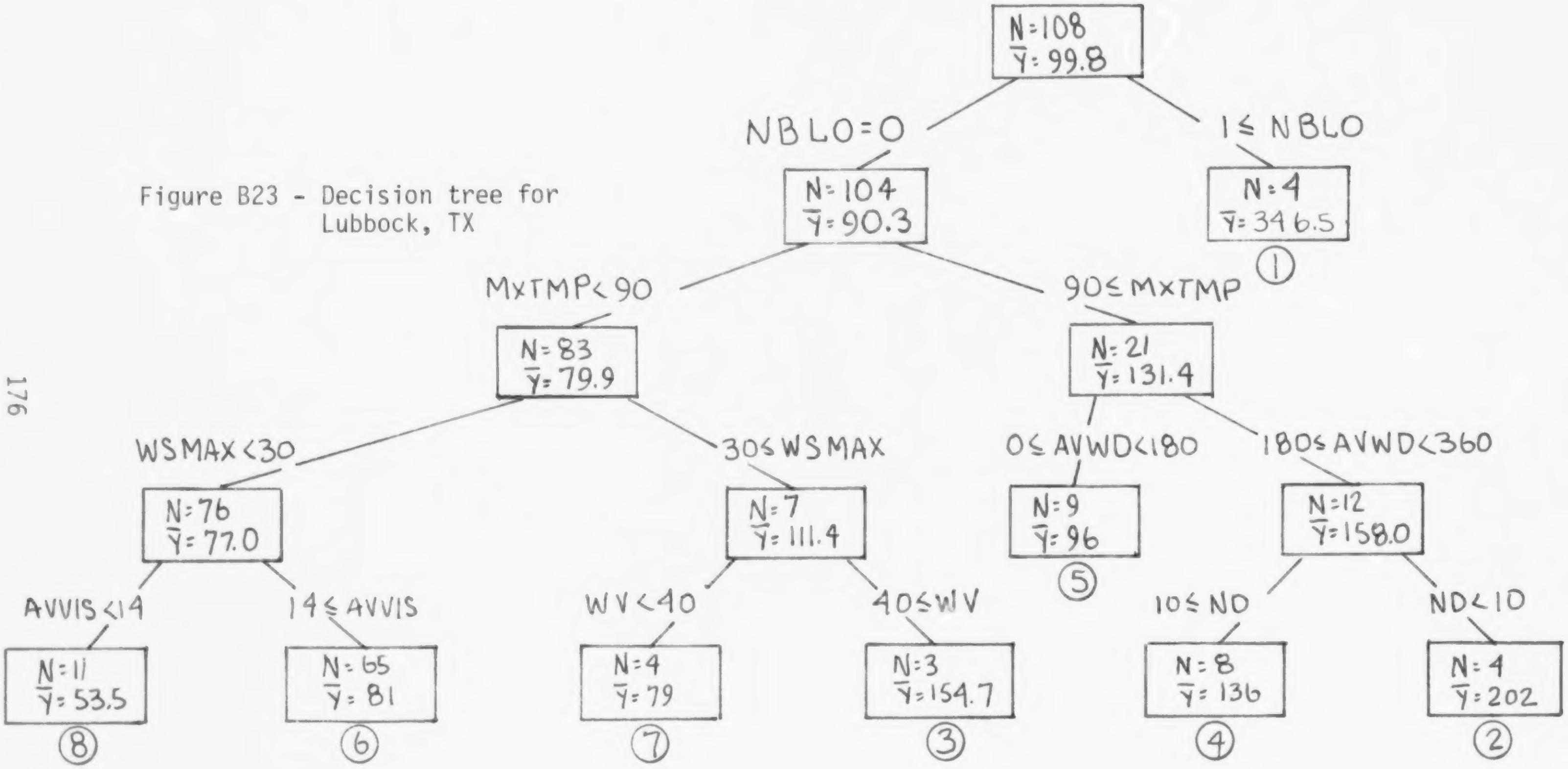

N=108
$\overline{Y}$=99.8
NBLO=0
1≤NBLO
N=104
$\overline{Y}$=90.3
N=4
$\overline{Y}$=346.5
①
MxTMP<90
90≤MxTMP
N=83
$\overline{Y}$=79.9
N=21
$\overline{Y}$=131.4
WSMAX<30
30≤WSMAX
0≤AVWD<180
180≤AVWD<360
N=76
$\overline{Y}$=77.0
N=7
$\overline{Y}$=111.4
N=9
$\overline{Y}$=96
⑤
N=12
$\overline{Y}$=158.0
AVVIS<14
14≤AVVIS
WV<40
40≤WV
10≤ND
ND<10
N=11
$\overline{Y}$=53.5
⑧
N=65
$\overline{Y}$=81
⑥
N=4
$\overline{Y}$=79
⑦
N=3
$\overline{Y}$=154.7
③
N=8
$\overline{Y}$=136
④
N=4
$\overline{Y}$=202
②

Figure B23 - Decision tree for
Lubbock, TX

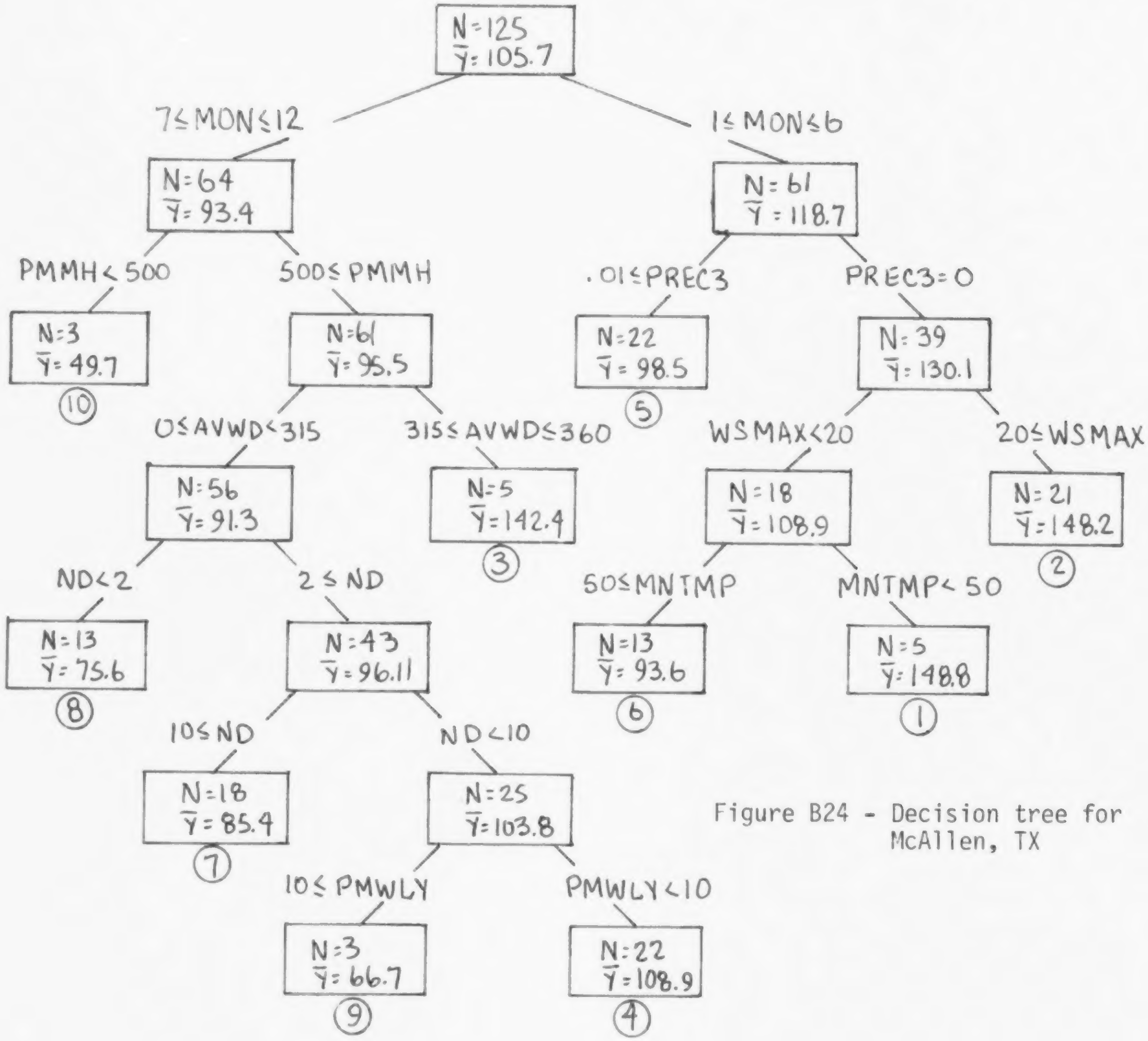

N=125
Ȳ=105.7
7≤MON≤12
1≤MON≤6
N=64
Ȳ=93.4
N=61
Ȳ=118.7
PMMH<500
500≤PMMH
.01≤PREC3
PREC3=0
N=3
Ȳ=49.7
⑩
N=61
Ȳ=95.5
N=22
Ȳ=98.5
⑤
N=39
Ȳ=130.1
0≤AVWD<315
315≤AVWD≤360
WSMAX<20
20≤WSMAX
N=56
Ȳ=91.3
N=5
Ȳ=142.4
③
N=18
Ȳ=108.9
N=21
Ȳ=148.2
②
ND<2
2≤ND
50≤MNTMP
MNTMP<50
N=13
Ȳ=75.6
⑧
N=43
Ȳ=96.11
N=13
Ȳ=93.6
⑥
N=5
Ȳ=148.8
①
10≤ND
ND<10
N=18
Ȳ=85.4
⑦
N=25
Ȳ=103.8
10≤PMWLY
PMWLY<10
N=3
Ȳ=66.7
⑨
N=22
Ȳ=108.9
④

Figure B24 - Decision tree for
McAllen, TX

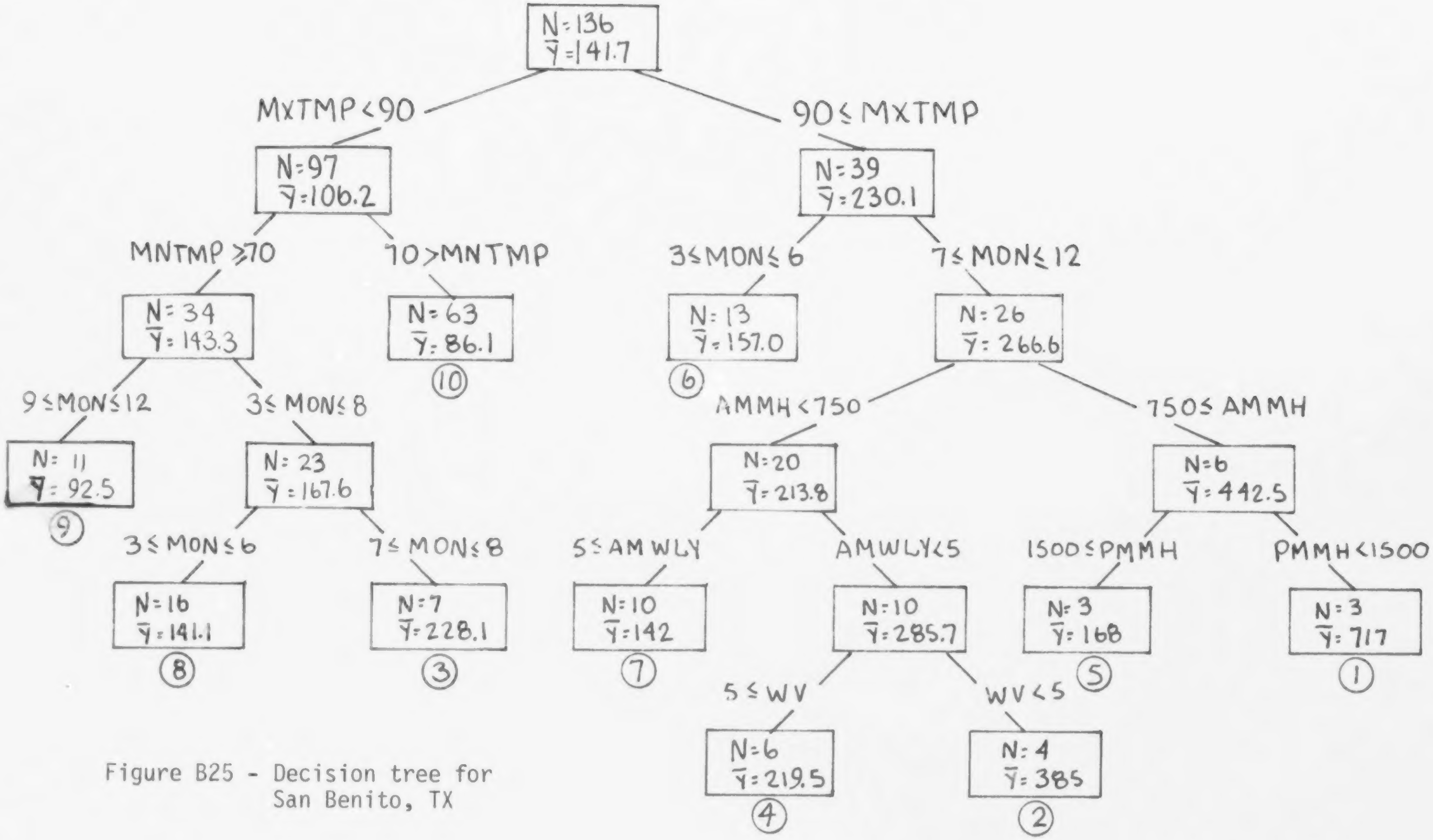

N=136
ȳ=141.7
MXTMP<90
90≤MXTMP
N=97
ȳ=106.2
N=39
ȳ=230.1
MNTMP≥70
70>MNTMP
3≤MON≤6
7≤MON≤12
N=34
ȳ=143.3
N=63
ȳ=86.1
⑩
N=13
ȳ=157.0
⑥
N=26
ȳ=266.6
9≤MON≤12
3≤MON≤8
AMMH<750
750≤AMMH
N=11
ȳ=92.5
⑨
N=23
ȳ=167.6
N=20
ȳ=213.8
N=6
ȳ=442.5
3≤MON≤6
7≤MON≤8
5≤AMWLY
AMWLY<5
1500≤PMMH
PMMH<1500
N=16
ȳ=141.1
⑧
N=7
ȳ=228.1
③
N=10
ȳ=142
⑦
N=10
ȳ=285.7
N=3
ȳ=168
⑤
N=3
ȳ=717
①
5≤WV
WV<5
N=6
ȳ=219.5
④
N=4
ȳ=385
②

Figure B25 - Decision tree for
San Benito, TX

APPENDIX C

INTERPRETATIONS OF AID DECISION TREES

SITE/CONCLUSION	CLASS	TERMINAL NODE #'s	AVERAGE TSP LEVEL ($\mu g/m^3$)	NUMBER OF DATA POINTS	QUALITATIVE DESCRIPTION
	All data	All 10	99.1	166	- - - - - - - - - - - - -
Forrest City, Arkansas Man-made sources (at least partly dust related) to the southwest or west. Also a possibility of contributions from wind-blown dust	I	1,2,3	156.1	34	high afternoon mixing height; very low wind speed or high afternoon mixing height; moderate to high wind speed; westerly wind; moderately dry ground; very low morning mixing height or high afternoon mixing height; moderate to high wind speed; southwesterly wind; moderately dry ground; moderate to high morning mixing height; warm night
	II	4	116.0	7	low afternoon mixing height; low morning mixing height
	III	5,6	102.3	29	high afternoon mixing height; moderate to high wind speed; southwesterly wind; moderately dry ground; very low morning mixing height; warm night or high afternoon mixing height; moderate to high wind speed; westerly wind; moderately dry ground; moderate to high morning mixing height; cool night
	IV	7,9	85.1	45	high afternoon mixing height; moderate to high wind speed; easterly wind or high afternoon mixing height; moderate to high wind speed; westerly wind; very wet ground
	V	8,10	69.2	51	low afternoon mixing height; moderate to high morning mixing height

Figure C1 - Interpretation of decision-tree for Forrest City, AR

SITE/CONCLUSION	CLASS	TERMINAL NODE #'s	AVERAGE TSP LEVEL ($\mu g/m^3$)	NUMBER OF DATA POINTS	QUALITATIVE DESCRIPTION
	All data	All 10	92.4	168	- - - - - - - - - - - -
Helena, AR A dominant man-made source to the southwest, probably dust related. Also, a possibility of contributions from wind-blown dust	I	1,3	172.7	15	southwesterly wind; moderately low RH; moderately high transport wind; moderately moist ground
	II	2,4,5,6	115.9	42	southwesterly wind; moderately low RH; varied transport winds and varied ground wetness
	III	8	83.2	11	southwesterly wind; high RH
	IV	7,9,10	70.9	100	wind not from southwest
Jonesboro, AR Man-made sources, some probably dust related. Also some minor contributions from wind-blown dust	All Data	All 10	80.3	165	- - - - - - - - - - - - -
	I	1,2	143.4	9	low wind speeds
	II	3,4	111.3	30	hot; low morning and high afternoon mixing heights; moderate wind speed
	III	5,6	89.8	30	cool; Nov. to Dec.; moderate wind speed or hot; low morning, and high afternoon moxing height; moderate to high wind speed
	IV	7,8	71.1	25	hot; moderate to high max wind speed; low morning and afternoon mixing height or hot; moderate to high max wind speed; high morning mixing height
	V	9,10	58.4	71	cool; Nov. to Dec.; high winds or cool; Jan. to Oct.; moderate to high winds

Figure C2 - Interpretation of decision-trees for Helena and Jonesboro, AR

SITE/CONCLUSION	CLASS	TERMINAL NODE #'s	AVERAGE TSP LEVEL ($\mu g/m3$)	NUMBER OF DATA POINTS	QUALITATIVE DESCRIPTION
	All data	All 10	78.4	143	- - - - - - - - - - - - -
Springdale, AR Man-made sources, mostly dust related. Also, some wind blown dust. Dust sources may be most intense during Jan. to Aug. (possibly agricultural) and in the westerly direction.	I	1,2,6	114.8	20	moderately dry ground; moderately high max wind speed; Jan. to Aug.; westerly wind; low visibility or moderately dry ground; low max wind speed; low afternoon mixing height
	II	3	109.4	10	very dry ground; high max wind speed; Jan. to Aug.
	III	4,5	84.7	37	moderately dry ground; moderately high max wind speed; Jan. to Aug.; westerly wind; high visibility; low RH or moderately dry ground; low max wind speed; high afternoon mixing height
	IV	7,8,9	66.1	50	moderately dry ground; moderately high max wind speed; Sept. to Dec. or moderately dry ground; moderately high max wind speed; Jan. to Aug.; easterly wind or moderately dry ground; moderately high max wind speed; Jan. to Aug.; westerly wind; good visibility; high RH
	V	10	53.1	26	Rain

Figure C3 - Interpretation of decision-tree for Springdale, AR

SITE/CONCLUSION	CLASS	TERMINAL NODE #'s	AVERAGE TSP LEVEL ($\mu g/m^3$)	NUMBER OF DATA POINTS	QUALITATIVE DESCRIPTION
	All data	All 10	104.3	168	- - - - - - - - - - - -
Stuttgart (#1), AR	I	1,2,3,6	168.3	25	northeasterly wind
Dominant man-made source toward the north or northeast, mostly dust related, probably a point source. Also some minor contributions from wind-blown dust.	II	4,5	148.0	16	north-northwesterly wind; moderately dry ground; moderately high max wind speed or wind not northeasterly; moderately dry ground; low max wind speed
	III	7,8	119.2	23	east around to northwesterly wind; moderately dry ground; moderately high max wind speed; moderately low wind variability or very high wind variability with very high aft. mixing height
	IV	9	85.5	61	east to northwesterly wind; moderately dry ground; moderately high max wind speed; very high wind variability; moderate to low afternoon mixing height
	V	10	69.4	43	wet ground
	All data	All 10	191.6	126	- - - - - - - - - - - - - -
Stuttgart (#2), AR	I	1	592.3	3	dry ground; low max wind speed
Man-made dust sources. Also, some contributions from wind-blown dust.	II	2,3	303.6	23	dry ground; high aft. mixing height; hot; very high wind speeds or moist ground; very high afternoon mixing height
	III	4,5	214.3	18	dry ground; moderate winds; high aft. mixing height; hot or dry ground; moderate to high max wind; high aft. mixing height; moderate to cool
	IV	6,7	175.4	27	slightly moist ground; moderate to low aft. mixing height; hot or dry ground; moderate to high max wind; low aft. mixing height
	V	8,9,10	122.1	55	very moist ground; moderate to low afternoon mixing height; hot or moist ground; moderate to low afternoon mixing height; cool

Figure C4 - Interpretation of decision-trees for Stuttgart (#1 and #2), AR

SITE/CONCLUSION	CLASS	TERMINAL NODE #'s	AVERAGE TSP LEVEL ($\mu g/m^3$)	NUMBER OF DATA POINTS	QUALITATIVE DESCRIPTION
Albuquerque (#1), NM Man-made sources (mostly dust related). Also, a major contribution from wind-blown dust.	All data	All 10	113.2	96	- - - - - - - - - - - -
	I	1	223.4	5	dry ground; low max wind speed; highly variable winds
	II	2	161.4	5	dry ground; very high max wind speed; Mar. to Dec.
	III	3,7	147.3	21	very dry ground; moderate to high max wind speed; low visibility
	IV	4,5,6,9	106.6	36	dry ground; low max wind speed; moderate to low wind variability or very dry ground; moderate to high max wind speed; good visibility or dry ground; moderate to high max wind speed; Mar. to Dec.
	V	8,10	69.9	29	moist ground or dry ground; moderate to high max wind speed; Jan. to Feb.

Figure C5 - Interpretation of decision-tree for Albuquerque (#1), NM

SITE/CONCLUSION	CLASS	TERMINAL NODE #'s	AVERAGE TSP LEVEL ($\mu g/m^3$)	NUMBER OF DATA POINTS	QUALITATIVE DESCRIPTION
Albuquerque (#2), NM Man-made sources, mostly dust related. Also, wind-blown dust. Wind-blown dust sources and/or man-made dust sources may be concentrated toward the southeast.	All data	All 10	80.8	95	- - - - - - - - - - - - -
	I	1,2	157.6	9	very low RH; low aft. mixing height or very low RH; high aft. mixing height; high transport winds; southeasterly winds
	II	3,4	112.1	11	moderate to high RH; low morning mixing height; winds from east round to north; moderate to low wind variability; very cold night or moderate to high RH; low morning mixing height; winds from east around to north; high wind variability
	III	5	99.4	8	very low TH; high aft. mixing height; high transport wind; westerly wind
	IV	6,7,10	72.7	32	very low RH; high aft. mixing height; low transport wind or moderate to high RH; low morning mixing height; winds from east around to north; moderate to low wind variability; not extremely cold night
	V	8,9	55.3	35	moderate to high RH, low morning mixing height; northeasterly wind or moderate to high RH; high morning mixing height

Figure C6 - Interpretation of decision-tree for Albuquerque (#2), NM

SITE/CONCLUSION	CLASS	TERMINAL NODE #'s	AVERAGE TSP LEVEL (µg/m3)	NUMBER OF DATA POINTS	QUALITATIVE DESCRIPTION
Dona Ana, NM	All data	All 10	98.6	135	- - - - - - - - - - - - -
Man-made dust sources, possibly concentrated toward the northeast. Possibly some minor contributions from wind-blown dust.	I	1,2,3,4	180.4	20	very low RH; good visibility; very low transport wind or very low RH; low visibility or very low RH; good visibility; moderate to high transport wind; high wind variability; northeasterly wind
	II	5,6	121.1	32	moderate to high RH; low morning mixing height; dry ground or very low RH; good visibility; moderate to high transport wind; high wind variability; not northeasterly wind
	III	7	96.9	24	very low RH; good visibility; moderate to high transport wind; low wind variability
	IV	8,9	65.7	50	moderate to high RH; low morning mixing height; possibly moist ground or moderate to high RH; high morning mixing height; light or no rain
	V	10	24.3	9	moderate to high RH; high morning mixing height; rain

Figure C7 - Interpretation of decision-tree for Dona Ana, NM

SITE/CONCLUSION	CLASS	TERMINAL NODE #'s	AVERAGE TSP LEVEL ($\mu g/m^3$)	NUMBER OF DATA POINTS	QUALITATIVE DESCRIPTION
Las Cruces, NM Man-made sources. Also, significant contributions from wind-blown dust. Dust sources may be most intense during winter and spring (possibly agriculture related).	All data	All 10	81.9	113	- - - - - - - - - - - -
	I	1,2,3	166.2	12	Nov. to Dec.; low average and high max wind speed or Jan. to June; moderate to low average wind speed; low morning mixing height; moderately dry ground or Jan. to June; very high average wind speed
	II	4,5,6,7	95.3	33	July to Oct.; low morning mixing height; low max wind speed or Jan. to June; moderate to low average wind speed; low morning mixing height; dry ground or Nov. to Dec.; high average wind speed or Nov. to Dec.; low wind speeds
	III	8,9	70.3	43	Jan. to June; moderate to low average wind speed; high morning mixing height or July to Oct.; low morning mixing height; high max wind speed
	IV	10	43.7	25	July to Oct.; high morning mixing height

Figure C8 - Interpretation of decision-tree for Las Cruces, NM

SITE/CONCLUSION	CLASS	TERMINAL NODE #'s	AVERAGE TSP LEVEL ($\mu g/m^3$)	NUMBER OF DATA POINTS	QUALITATIVE DESCRIPTION
Raton, NM Man-made sources, mostly to the east. Also a possibility of wind-blown dust. Dust sources may be most intense in spring and least intense in winter (possibly agriculture related).	All data	All 10	50.2	85	- - - - - - - - - - - -
	I	1	106.3	3	Mar. to Dec.; moderate to high temperature; easterly wind; low average wind speed
	II	2,3	71.4	14	Mar. to Apr.; moderate to high temperature; westerly wind or Mar. to Dec.; moderate to high temperature; southeasterly wind; high wind speeds
	III	4	53.4	20	Mar. to Dec.; moderate to high temperature; north to southeasterly wind; high wind speeds
	IV	5,6	46.5	30	May to Dec.; moderate to high temperature; westerly wind; low Rh
	V	7,8,9,10	27.1	18	Mar. to Dec.; moderate to high temperature; easterly wind; moderate wind speeds or Jan. to Feb. or May to Dec.; moderate to high temperature; westerly wind; high RH or Mar. to Dec.; cold

Figure C9 - Interpretation of decision-tree for Raton, NM

SITE/CONCLUSIONS	CLASS	TERMINAL NODE #'s	AVERAGE TSP LEVEL ($\mu g/m^3$)	NUMBER OF DATA POINTS	QUALITATIVE DESCRIPTION
Bethany, OK	All data	All 10	82.7	186	- - - - - - - - - - - -
Major contributions from wind-blown dust. Also, some dust related man-made sources.	I	1,2	142.4	15	low RH or moderate to high RH; moderate to high temperature; dry ground; very high wind speed
	II	4,5	98.2	46	moderate RH; moderate to high temperature; dry ground; moderate to low wind speed
	III	3,6	94.0	27	moderate to high RH; moderate to high temperature; moist ground; low transport wind
	IV	7,8,9,10	63.1	98	moderate to high RH; moderate to high temperature; moist ground; high transport wind or high RH; moderate to high temperature; dry ground; moderate to low wind speed or moderate to high RH; cold

Figure C10 - Interpretation of decision-tree for Bethany, OK

SITE/CONCLUSIONS	CLASS	TERMINAL NODE #'s	AVERAGE TSP LEVEL ($\mu g/m^3$)	NUMBER OF DATA POINTS	QUALITATIVE DESCRIPTION
Muskogee, OK	All data	All 10	60.2	217	- - - - - - - - - - - -
Man-made dust sources, most intense from January to August (possibly agriculture related). Also a possibility of contributions from wind-blown dust.	I	1,2	102.2	9	moderate to high temp.; very dry ground or moderate to high temp.; Jan. to Aug.; moderately dry ground; very high wind variability
	II	3,4,5	77.8	53	moderate to high temp.; Jan. to Aug.; moderately dry ground; moderate to low wind variability, morning mixing height, and transport wind speed or cool, low RH or cool; moderate to high RH; low wind speed
	III	6,7,8	58.4	54	moderate to high temp.; Jan. to Aug.; moist ground or moderate to high temp.; Jan. to Aug.; moderately dry ground; moderately low wind variability; high morning mixing height or moderate to high temp.; Jan. to Aug.; moderately dry ground; moderately low wind variability; moderately low morning mixing height; high transport winds
	IV	9	54.4	38	moderate to high temp.; possibly moist ground; Sept. to Dec.
	V	10	44.6	63	cool; moderate to high RH; high max wind speed

Figure C11 - Interpretation of decision-tree for Muskogee, OK

SITE/CONCLUSIONS	CLASS	TERMINAL NODE #'s	AVERAGE TSP LEVEL ($\mu g/m^3$)	NUMBER OF DATA POINTS	QUALITATIVE DESCRIPTION
Oklahoma City, OK	All data	All 10	46.3	137	- - - - - - - - - - - -
Man-made sources, mostly dust related. Sources are most intense toward the north and during the first half of the year (possibly agriculture related). Also a possibility of wind-blown dust.	I	1,2	123.9	9	very low RH; low visibility or moderate to high RH; north-northwesterly winds; dry ground
	II	3,4,5,6	59.3	27	moderate to high RH; mortheasterly wind; hot; Jan. to June or very low RH; good visibility or moderate RH; east around to northwesterly winds; hot; Jan. to June
	III	7,8	40.3	57	high RH; east around to northwesterly winds; hot; Jan. to June or moderate to high RH; not north-northwesterly winds; hot; July to Dec.
	IV	9,10	30.4	44	moderate to high RH; nor north-northwesterly wind; cool or moderate to high RH; northwesterly wind; rain

Figure C12 - Interpretation of decision-tree for Oklahoma City, OK

SITE/CONCLUSIONS	CLASS	TERMINAL NODE #'s	AVERAGE TSP LEVEL ($\mu g/m^3$)	NUMBER OF DATA POINTS	QUALITATIVE DESCRIPTION
Roger Mills, OK	All data	All 9	83.4	106	- - - - - - - - - - - -
Man-made sources, mostly dust related. Also a possibility of major contributions from wind-blown dust	I	1,2	188.2	10	very high nighttime temperature; low RH or moderate to cool night; very low RH
	II	4,5	107.6	10	moderate to cool night; moderate to high RH; very high wind variability or moderate to cool night; moderate to high RH; moderate to low wind variability; high aft. mixing height; very dry ground
	III	3,6	93.1	7	moderate to cool night; moderate to high RH; moderate to low wind variability; high aft. mixing height; heavy storm
	IV	7,8	68.6	68	moderate to cool night; moderate to high RH; moderate to low wind variability; high aft. mixing height; possibly moist ground or very high nighttime temperature; high RH
	V	9,10	49.8	11	moderate to cool night; moderate to high RH; moderate to low wind variability; low afternoon mixing height

Figure C13 - Interpretation of decision-tree for Roger Mills, OK

SITE/CONCLUSIONS	CLASS	TERMINAL NODE #'s	AVERAGE TSP LEVEL (μg/m3)	NUMBER OF DATA POINTS	QUALITATIVE DESCRIPTION
Sequoyah Co., OK	All data	All 10	49.2	119	- - - - - - - - - - - -
Man-made sources, mostly dust related. Also some minor contributions from wind-blown dust.	I	1,2	111.6	13	very hot; low RH; moderate to low max wind speed or moderate temperatures; moderately dry ground; low transport wind speed
	II	3,4,5	64.3	25	very hot; low RH; high max wind speed or moderate temperatures; moderately dry ground; moderate to high transport wind speed; high aft. mixing height or very hot; high RH
	III	6,7,8	45.7	33	moderate to cool temperatures; moderately dry ground; low transport wind speed or moderate temperatures; moderately dry ground; low transport wind; low afternoon mixing height
	IV	9,10	26.8	48	moderate to cool temperatures; moderately dry ground; high wind speed or moderate to cool temperature; wet ground

Figure C14 - Interpretation of decision-tree for Sequoyah Co., OK

SITE/CONCLUSION	CLASS	TERMINAL NODE #'s	AVERAGE TSP LEVEL ($\mu g/m^3$)	NUMBER OF DATA POINTS	QUALITATIVE DESCRIPTION
Tulsa (#1), OK Major contributions from wind-blown dust. Also some contributions from man-made dust sources.	All data	All 10	80.0	210	- - - - - - - - - - - -
	I	1,3	217.8	9	very high wind speed
	II	2,4	144.6	19	moderately high wind speed; warm; dry ground; low wind variability
	III	5,6,8	85.0	56	moderate to low wind speed; warm; dry ground; moderate to high wind variability or low wind speed; warm; dry ground; low wind variability or moderate to low wind speed; warm; moderately dry ground
	IV	7	80.6	17	low wind speed; moist ground
	V	9,10	54.7	109	moderate to low wind speeds; dry ground; cool or moderate wind speeds; moist ground

Figure C15 - Interpretation of decision-tree for Tulsa (#1), NM

SITE/CONCLUSIONS	CLASS	TERMINAL NODE #'s	AVERAGE TSP LEVEL ($\mu g/m^3$)	NUMBER OF DATA POINTS	QUALITATIVE DESCRIPTION
Tulsa (#2), OK Significant contributions from wind-blown dust. Also, man-made dust sources which may be most intense toward the south or west.	All data	All 10	76.1	232	- - - - - - - - - - - -
	I	1,3	140.8	10	very high wind speed
	II	2	114.0	22	moderate to low wind speed; moderately dry ground; high wind variability; moderate to low morning mixing height
	III	5	78.0	90	moderate to low wind speed; moderately dry ground; low wind variability
	IV	4,6,7	73.4	39	moderate to low wind speed; moderately moist ground; winds not from north-northeast or moderate to low wind speed; moderately dry ground; high wind variability; high morning mixing height
	V	8,9,10	54.2	71	moderate to low wind speed; wet ground; winds not from north-northeast or moderate to low wind speed; moderately moist ground; north-northeasterly wind

Figure C16 - Interpretation of decision-tree for Tulsa (#2), NM

SITE/CONCLUSION	CLASS	TERMINAL NODE #'s	AVERAGE TSP LEVEL ($\mu g/m^3$)	NUMBER OF DATA POINTS	QUALITATIVE DESCRIPTION
Corpus Christi (#1), TX Man-made sources, some dust related. Possibly some minor contributions from wind-blown dust.	All data	All 10	82.7	128	- - - - - - - - - - - - -
	I	1,2	146.3	7	low afternoon mixing height; very dry ground or moderate to low afternoon mixing height; possibly moist ground; moderate wind speeds
	II	3,4,7	90.8	46	moderate to low afternoon mixing height; possibly moist ground; moderate wind speeds
	III	5,6	82.4	39	moderate afternoon mixing height; very dry ground or moderate to low afternoon mixing height; possibly moist ground; high max wind speed; moderate to low RH; warm
	IV	8,9,10	60.2	36	high afternoon mixing height or moderate to low afternoon mixing height; possibly moist ground; high wind speed; moderate to low RH; cool or moderate to low afternoon mixing height; possibly wet ground; high wind speed; very high RH

Figure C17 - Interpretation of decision-tree for Corpus Christi (#1), TX

SITE/CONCLUSION	CLASS	TERMINAL NODE #'s	AVERAGE TSP LEVEL ($\mu g/m^3$)	NUMBER OF DATA POINTS	QUALITATIVE DESCRIPTION
Corpus Christi (#2), TX	All data	All 10	109.4	121	- - - - - - - - - - - - -
Probably a mixture of man-made sources and wind-blown dust. Sources tend to be strongest toward the north or east.	I	1	280.8	4	moderate to high mixing height; moderately cool nights
	II	2,3	147.5	16	moderate to high afternoon mixing height; moderate to warm night; low wind variability; north to southeasterly winds or high mixing heights; moderately cool nights
	III	4,5,6	122.2	33	moderate to high afternoon mixing height; moderate to warm night; low wind variability; winds from southeast arount to north; high transport wind or moderate to high afternoon mixing height; moderate to warm night; moderate to high wind variability; high transport wind or moderate mixing heights; moderately cool night
	IV	7,8	90.3	53	moderate to high afternoon mixing height; moderate to warm night; low wind variability; winds from southeast around to north; low transport wind or moderate to high afternoon mixing height; moderate to warm night; high wind variability; low transport wind
	V	9,10	62.6	15	moderate to high afternoon mixing height; cold night or low afternoon mixing height

Figure C18 - Interpretation of decision-tree for Corpus Christi (#2), TX

SITE/CONCLUSION	CLASS	TERMINAL NODE #'s	AVERAGE TSP LEVEL ($\mu g/m^3$)	NUMBER OF DATA POINTS	QUALITATIVE DESCRIPTION
Corpus Christi (#3), TX Wind-blown dust and man-made sources (both dust related and conventional).	All data	All 10	74.5	106	- - - - - - - - - - - - -
	I	1,2,4	123.0	10	very dry ground or not extremely dry ground; low visibility; high transport wind
	II	3,5,6	83.0	28	possibly moist ground; good visibility; low morning mixing height or not extremely dry ground; low visibility; low transport wind
	III	7,8	68.3	51	not extremely dry ground; high mixing heights; low transport wind or moderately dry ground; good visibility; low morning mixing height
	IV	9,10	50.5	17	not extremely dry ground; good visibility; high mixing heights with high wind or moderate mixing heights with little mixing height change during the day

Figure C19 - Interpretation of decision-tree for Corpus Christi (#3), TX

SITE/CONCLUSION	CLASS	TERMINAL NODE #'s	AVERAGE TSP LEVEL ($\mu g/m^3$)	NUMBER OF DATA POINTS	QUALITATIVE DESCRIPTION
Harlingen, TX Dust sources, both wind-blown and man-made.	All data	All 10	99.2	149	- - - - - - - - - - - -
	I	1	238.8	5	dry ground; warm night; high wind speed; high afternoon mixing height
	II	2,3	161.6	17	dry ground; warm night; high wind speed; moderate to low afternoon mixing height
	III	4,5,6,7	107.6	65	dry ground; moderate to cool night; low RH or very dry ground; moderate to cool night; high RH; or dry ground; warm night; moderate to low wind speed or moist ground; hot; moderate to high wind speed
	IV	8,9,10	62.1	62	moderately dry ground; moderate to cool night; high RH or moist ground; hot; low wind speed or moist ground; cool
Lubbock, TX Substantial contributions from wind-blown dust (blowing dust may not be noted by the airport observer on many occasions when it occurs). Significant dust sources (wind-blown or man-made) toward the west.	All data	All 8	99.8	108	- - - - - - - - - - - -
	I	1	346.5	4	blowing dust noted
	II	2,4	158.0	12	blowing dust not noted; hot; westerly wind
	III	3,7	111.4	7	blowing dust not noted; moderate to cool temperature; very high max wind speed
	IV	5,6,8	79.0	85	blowing dust not noted; hot; easterly wind or blowing dust not noted; moderate to cool temperatures; moderate to low max wind speed

Figure C20 - Interpretation of decision-trees for Harlingen and Lubbock, TX

SITE/CONCLUSION	CLASS	TERMINAL NODE #'s	AVERAGE TSP LEVEL ($\mu g/m^3$)	NUMBER OF DATA POINTS	QUALITATIVE DESCRIPTION
McAllen, TX	All data	All 10	105.7	125	- - - - - - - - - - - - -
Three source types apparent: 1. general man-made dust sources 2. wind-blown dust 3. man-made dust source to the north-northwest Sources may be intensified during first half of year (possibly agriculture related).	I	1,2,3	147.4	31	Jan. to June; dry ground; moderate to low max wind speed; cool night or Jan. to June; dry ground; very high max wind speed or July to Dec.; moderately high afternoon mixing height; north-northwesterly wind
	II	4	108.9	22	July to Dec.; moderately high afternoon mixing height; not north-northwesterly wind; moderately dry ground; moderate to low transport wind
	III	5,6	96.7	35	Jan. to June; moist ground or Jan. to June; dry ground; moderate to low max wind speed; warm night
	IV	7,8	81.3	31	July to Dec.; moderately high afternoon mixing height; not north-northwesterly wind; varied ground moisture
	V	9,10	58.2	6	July to Dec.; moderately high afternoon mixing height; not north-northwesterly wind; moderately dry ground; high transport wind or July to Dec.; low afternoon mixing height

Figure C21 - Interpretation of decision-tree for McAllen, TX

SITE/CONCLUSION	CLASS	TERMINAL NODE #'s	AVERAGE TSP LEVEL ($\mu g/m^3$)	NUMBER OF DATA POINTS	QUALITATIVE DESCRIPTION
San Benito, TX Man-made sources, probably dust related and most prevalent during warm months (possibly agriculture related). Also, a possibility of contributions from wind-blown dust.	All data	All 10	141.7	136	- - - - - - - - - - - -
	I	1,2,3,4	544.2	20	hot; July to Dec.; high morning and low afternoon mixing height or hot; July to Dec.; low morning mixing height; low transport wind or warm; July to August
	II	6,8	148.3	29	hot or warm; March to June
	III	5,7	148.0	13	hot; July to Dec.; high mixing heights or warm; July to Dec.; moderate to low morning mixing height; high transport wind
	IV	9,10	87.1	74	warm; September to December or moderate to cool

Figure C22 - Interpretation of decision-tree for San Benito, TX

APPENDIX D

TIME SERIES PLOTS OF TSP DATA

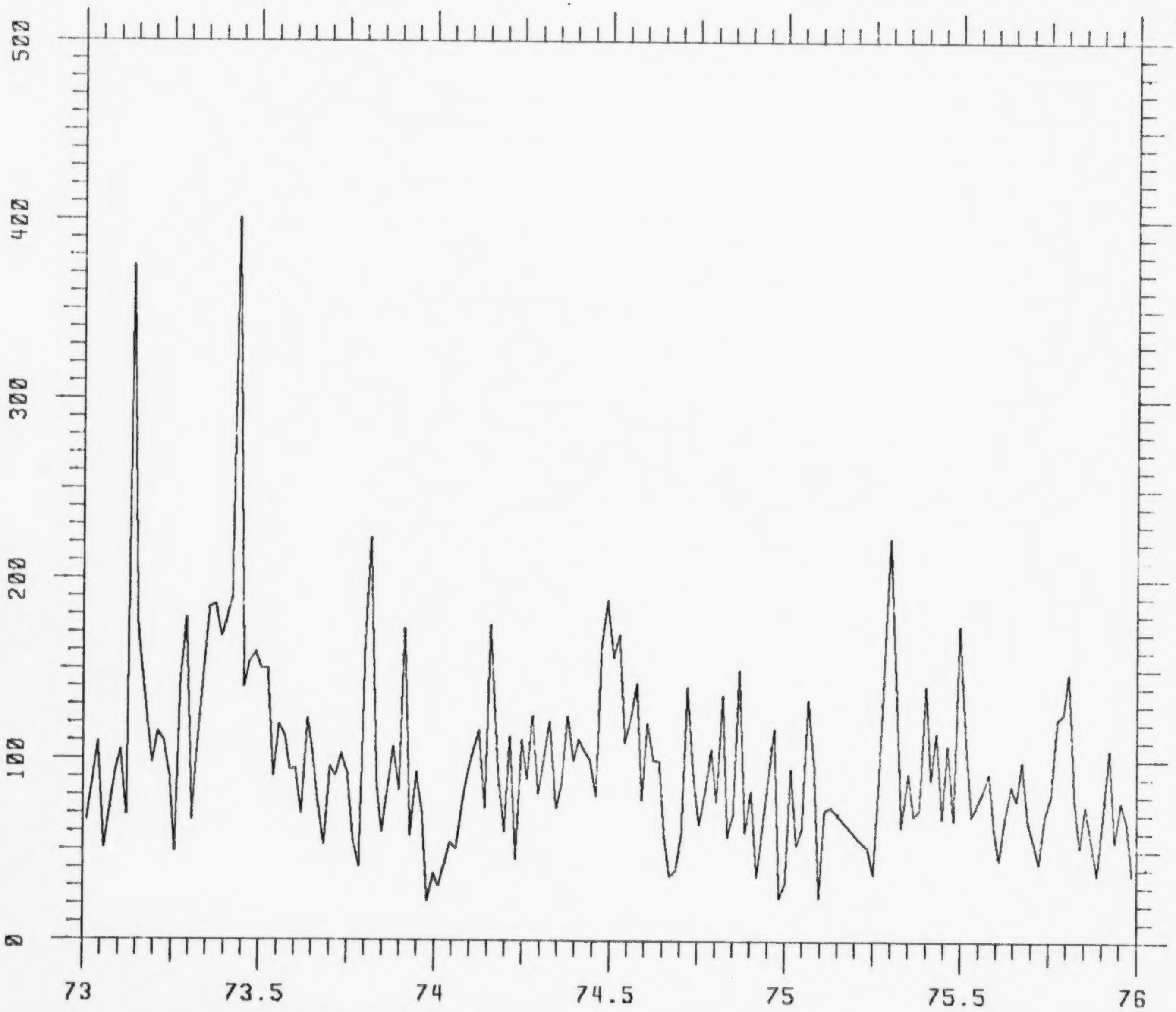


Figure D1 - Time series plot for Forrest City, AR

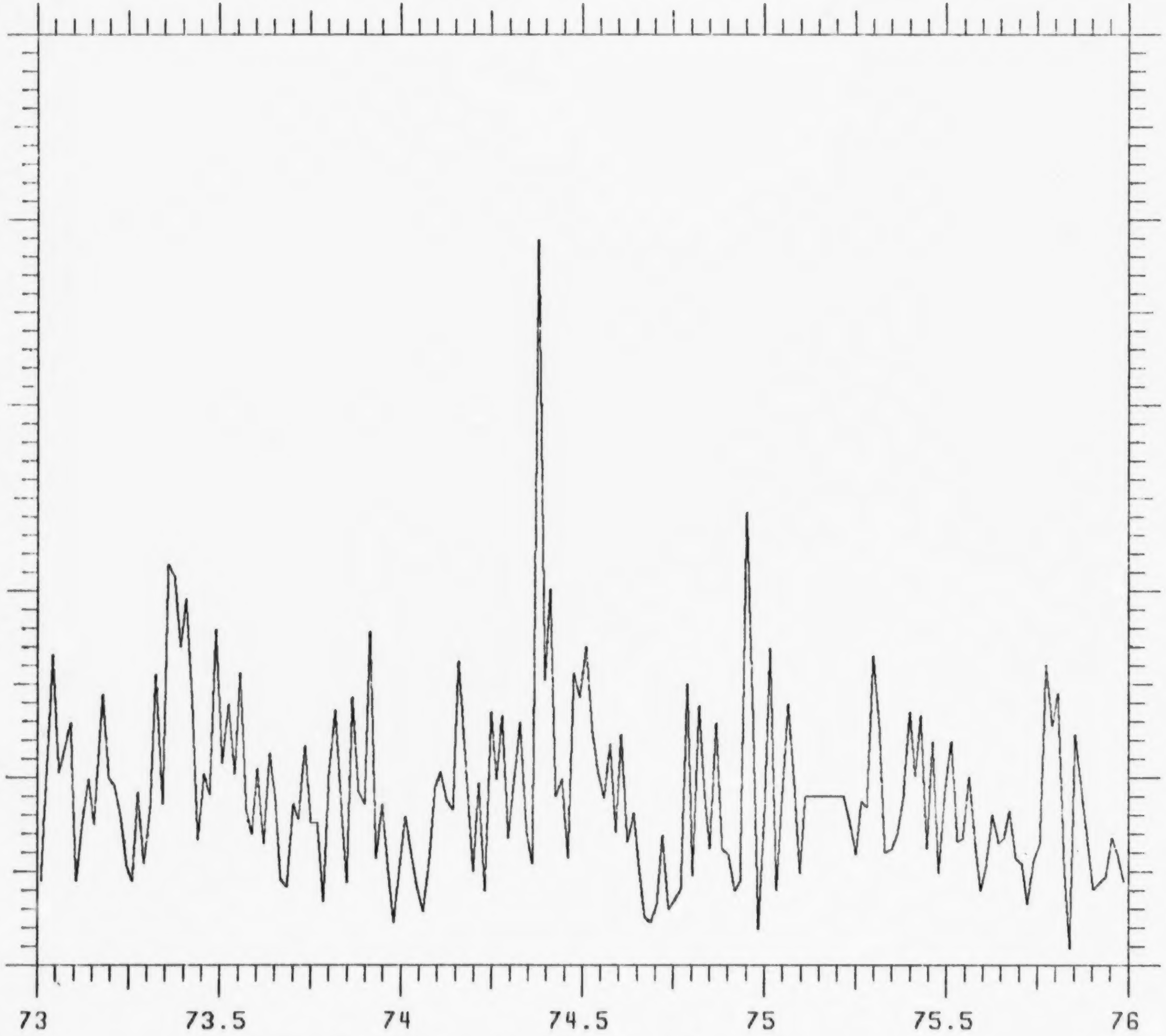


Figure D2 - Time series plot for Helena, AR

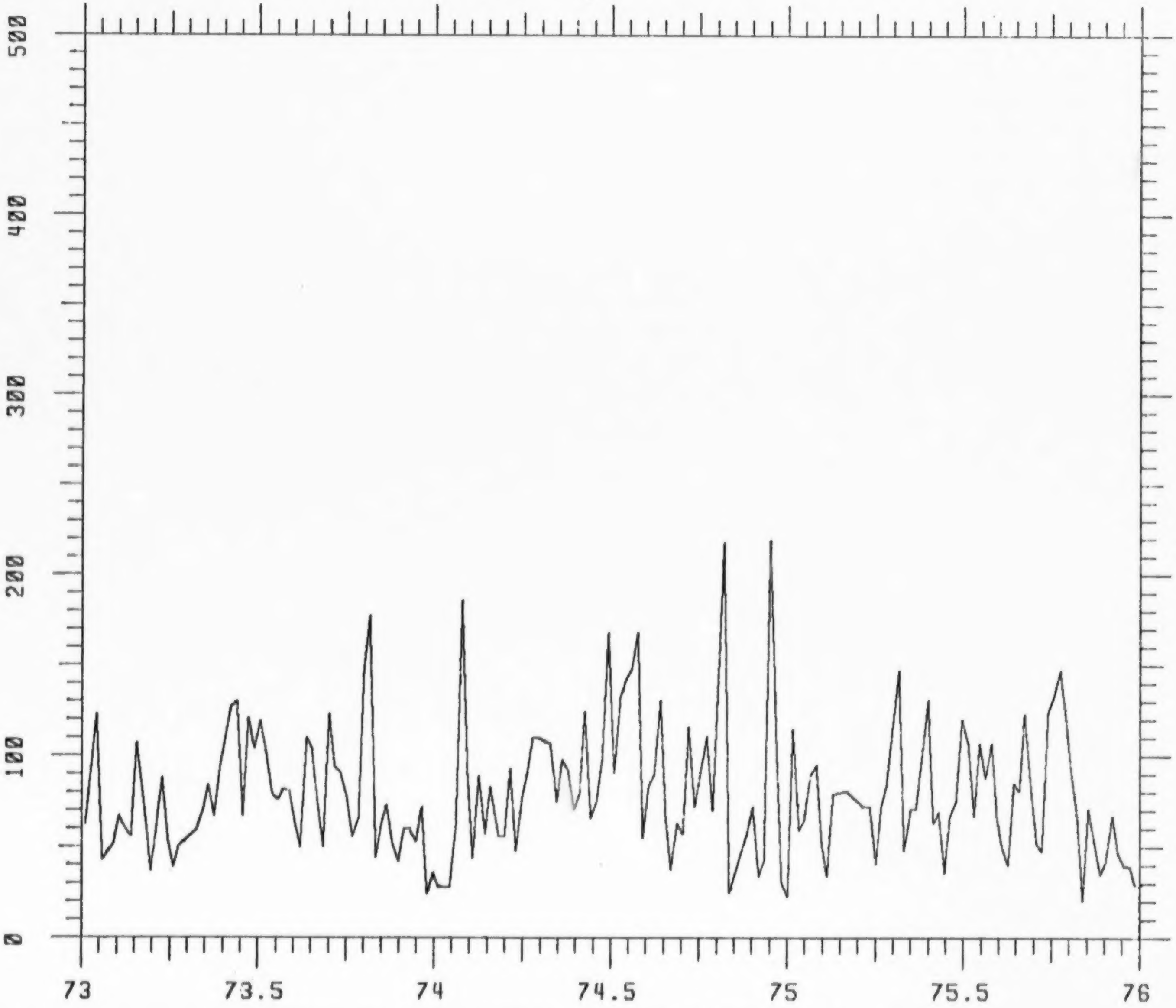


Figure D3 - Time series plot for Jonesboro, AR

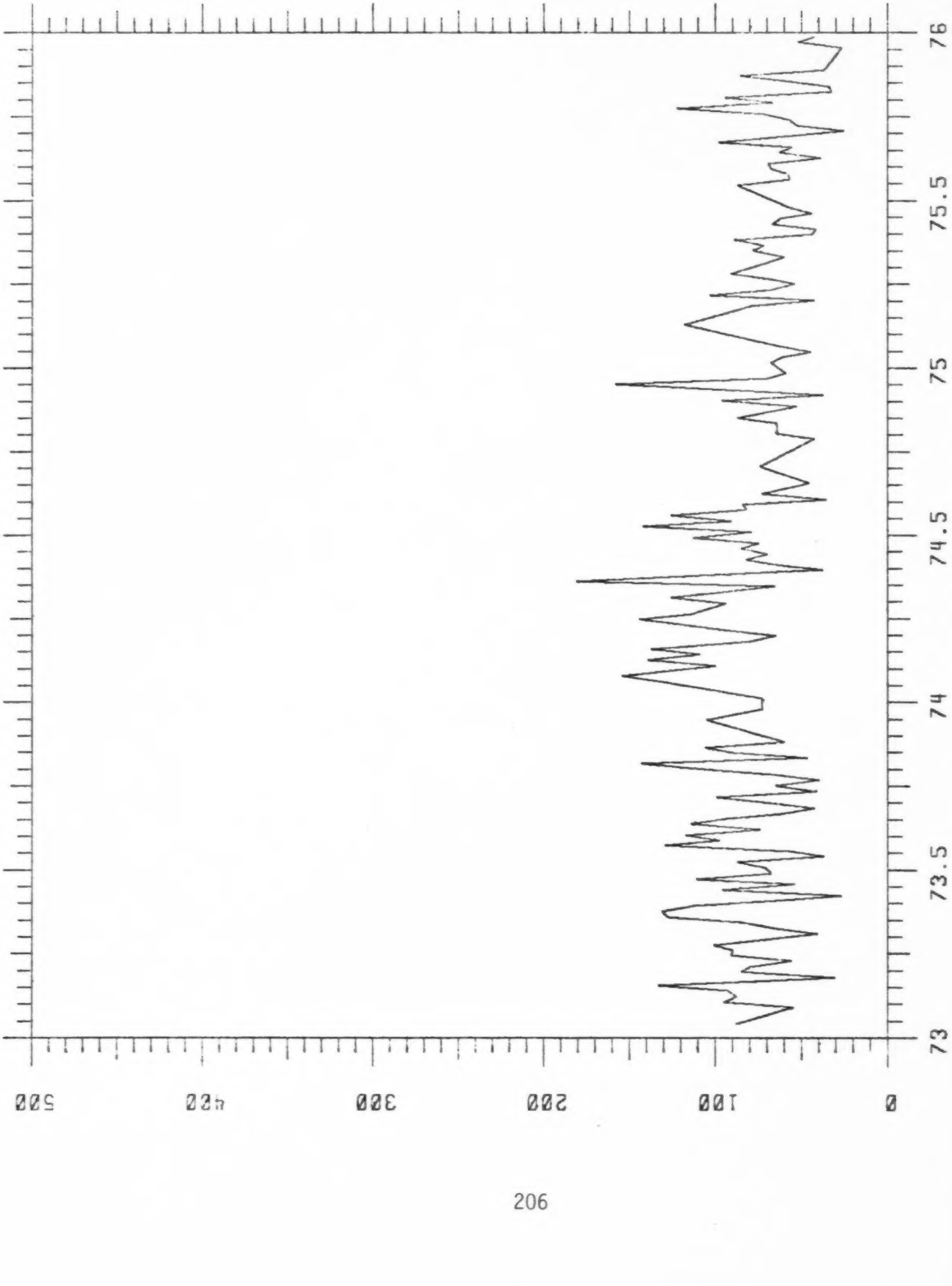


Figure D4 - Time series plot for Springdale, AR

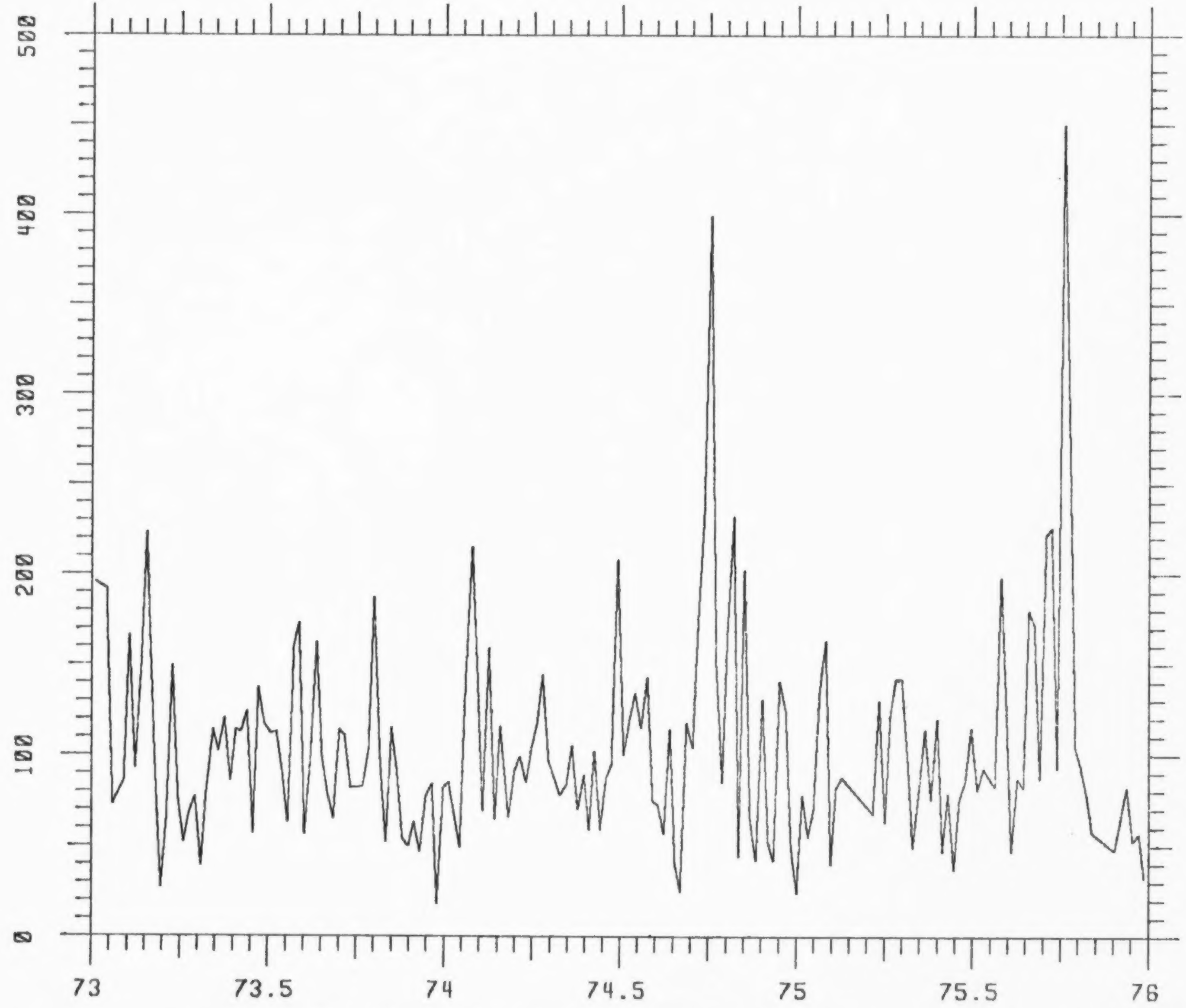


Figure D5 - Time series plot for Stuttgart (#1), AR

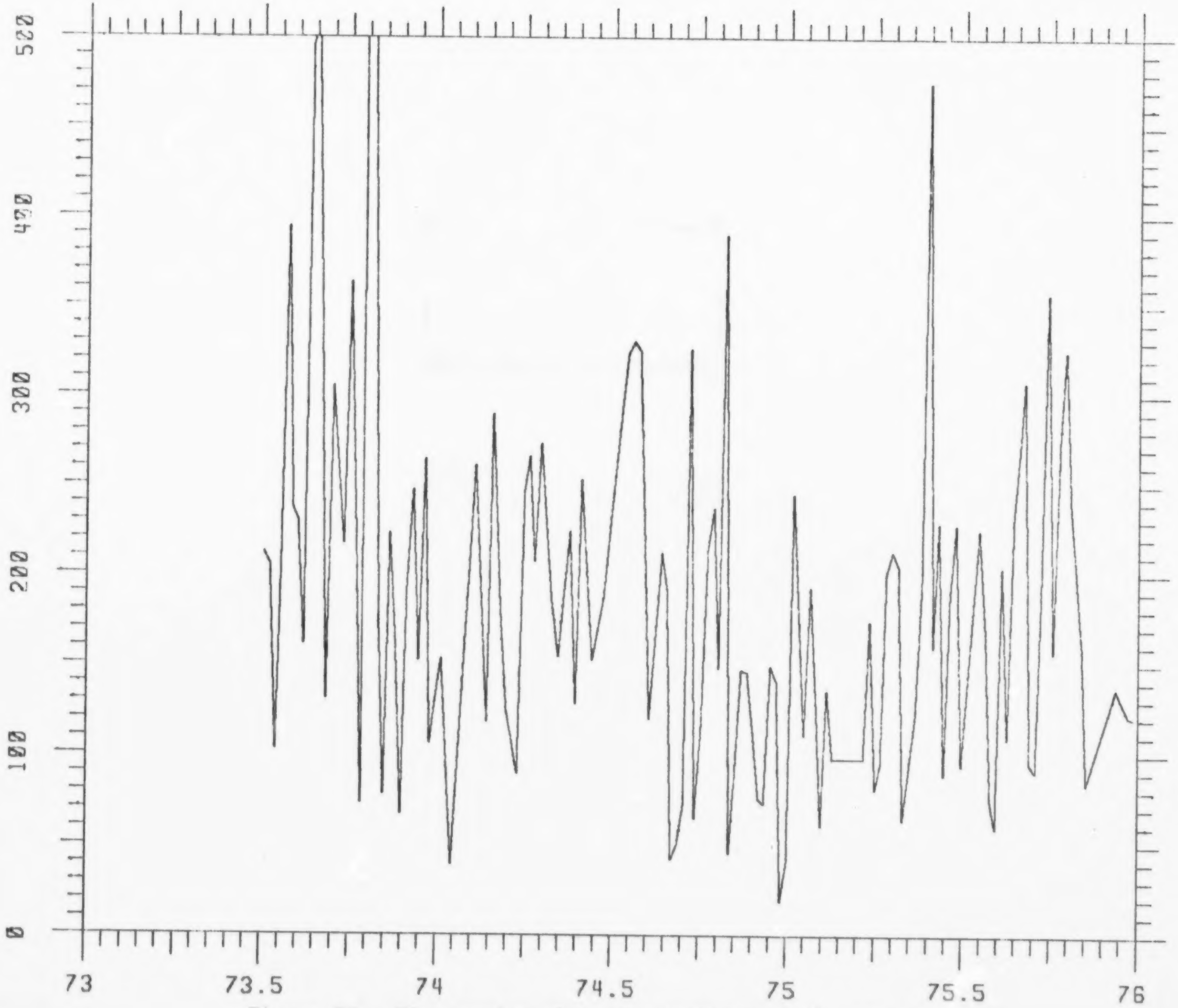


Figure D6 - Time series plot for Stuttgart (#2), AR

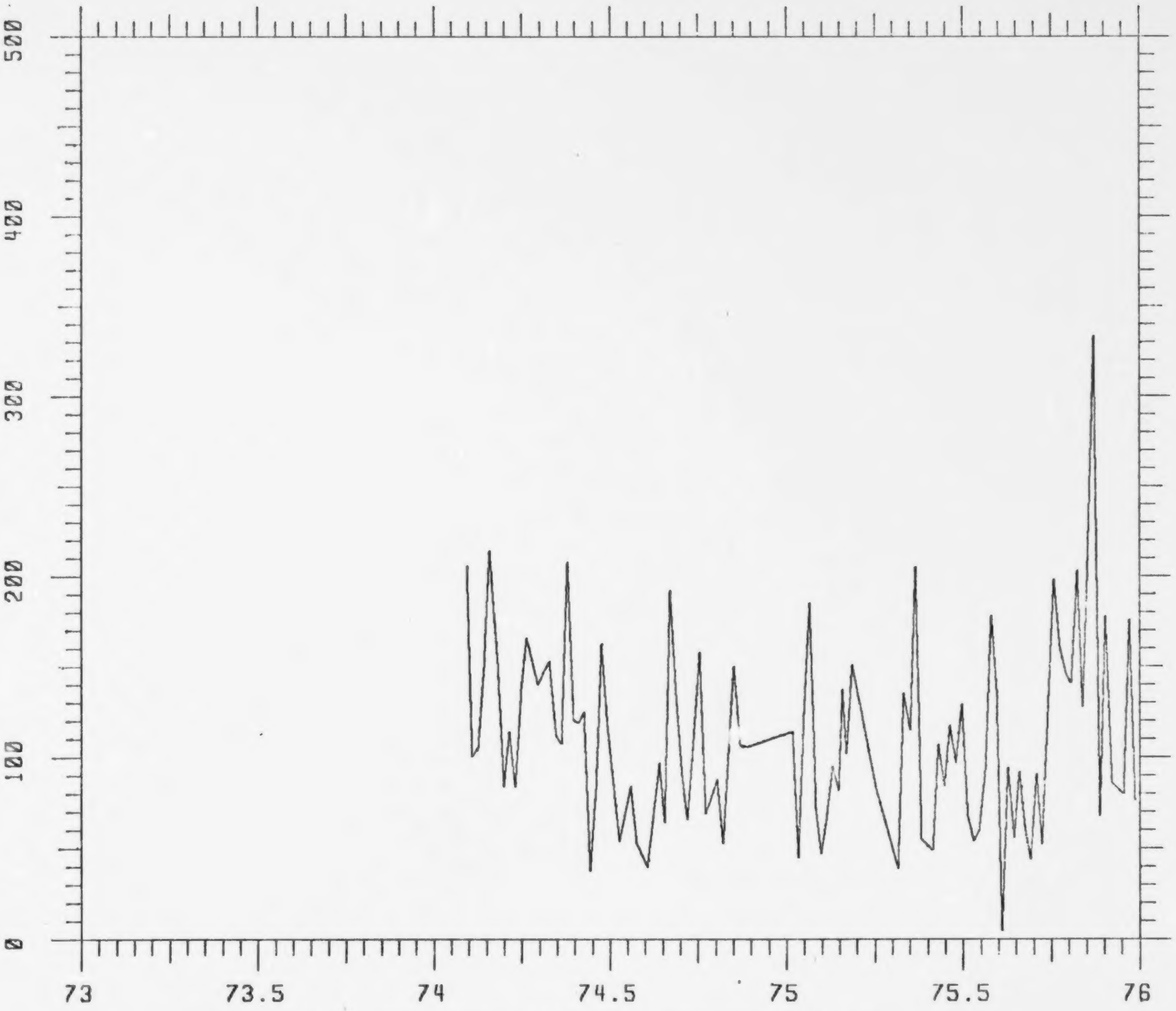


Figure D7 - Time series plot for Albuquerque (#1), NM

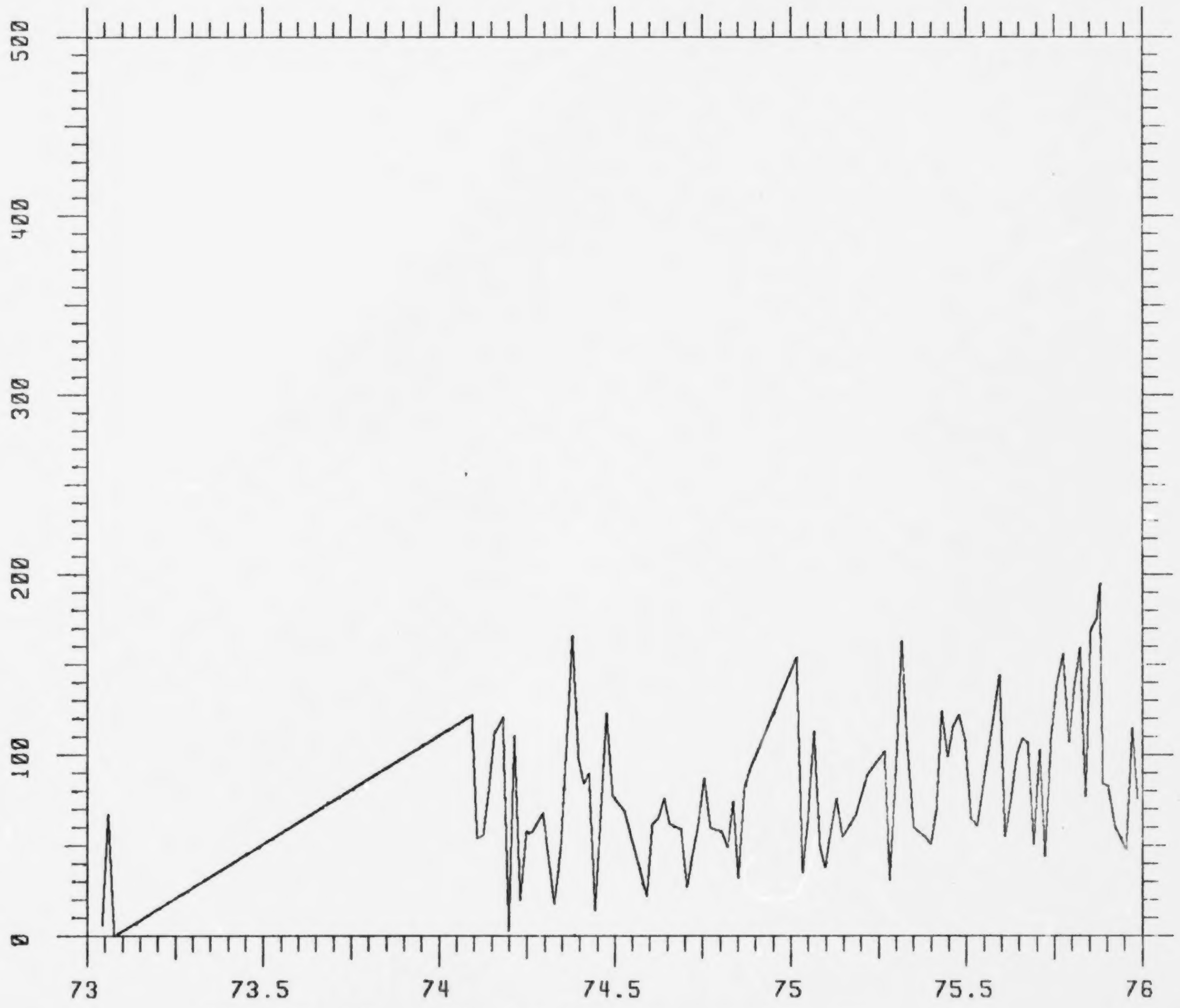


Figure D8 - Time series plot for Albuquerque (#2), NM

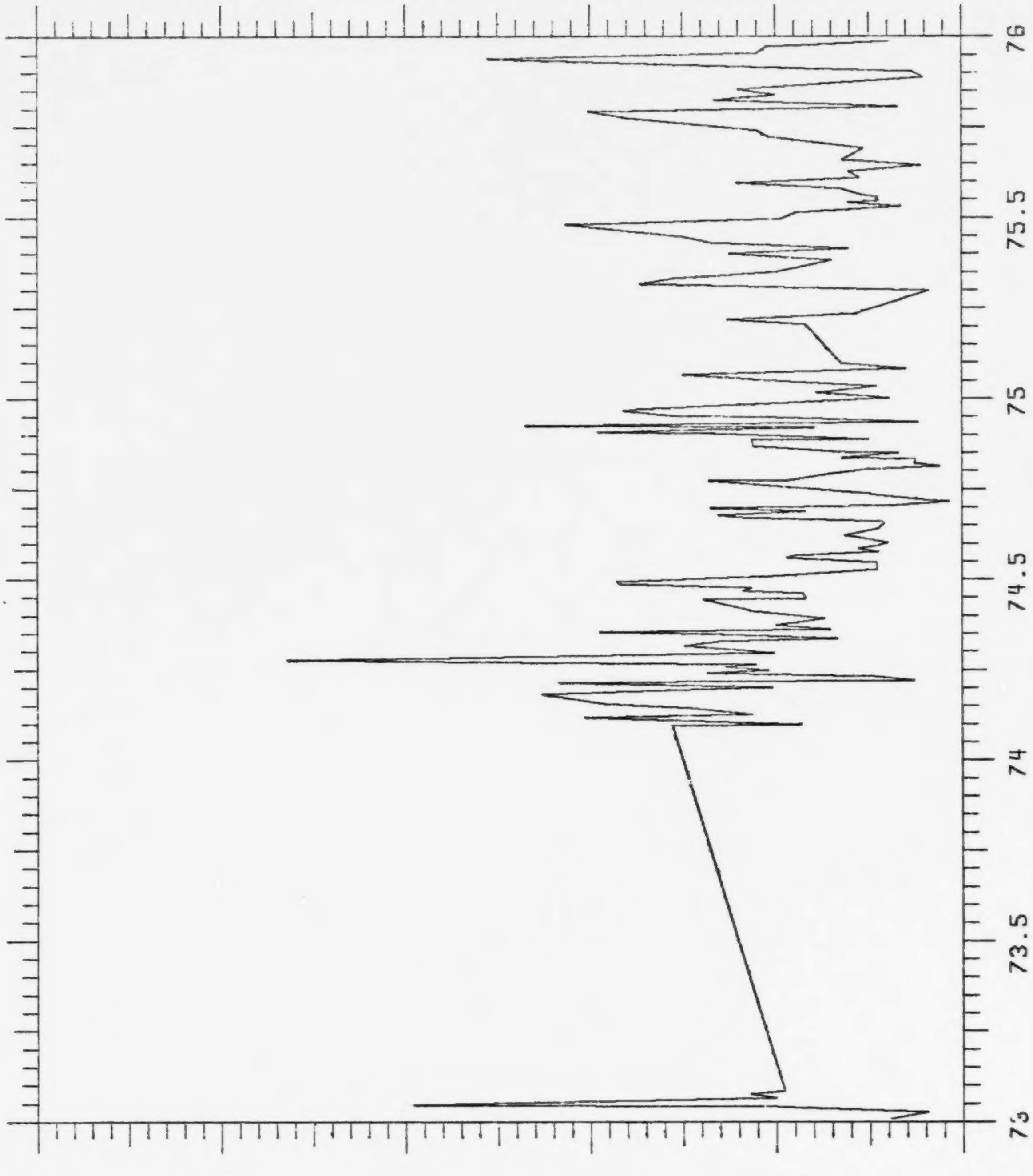


Figure D9 - Time series plot for Dona Ana, NM

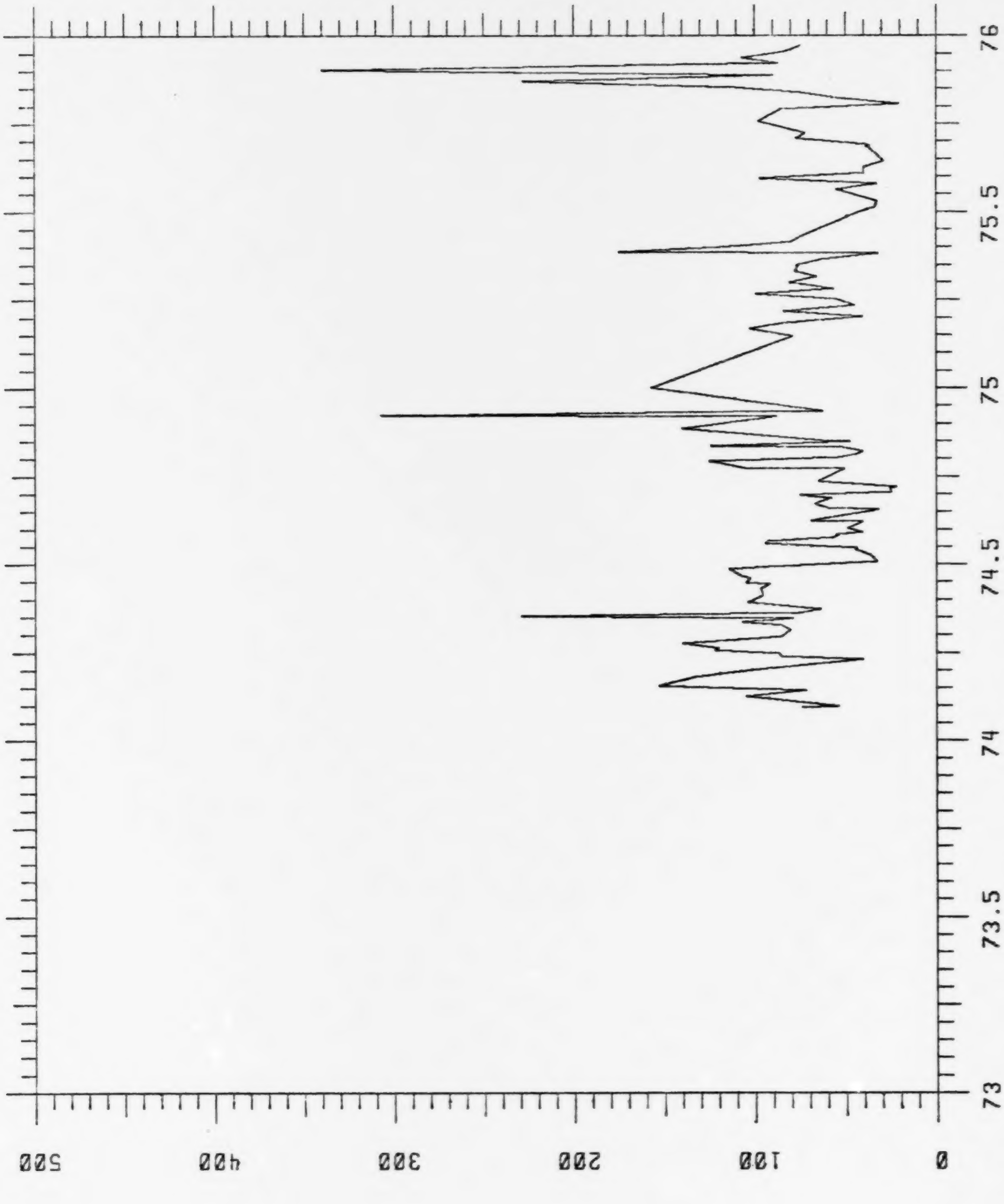


Figure D10 - Time series plot for Las Cruces, NM

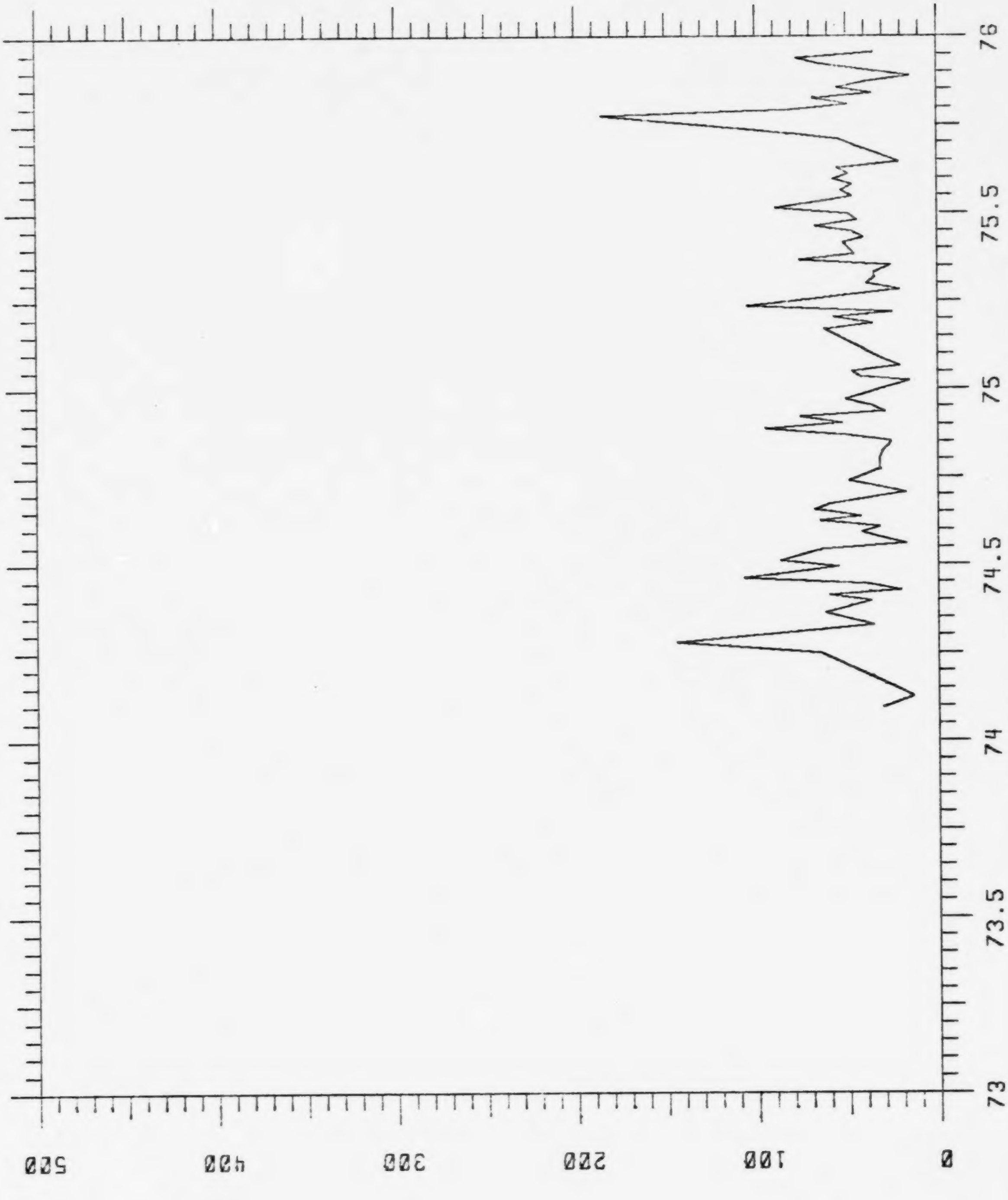


Figure D11 - Time series plot for Raton, NM

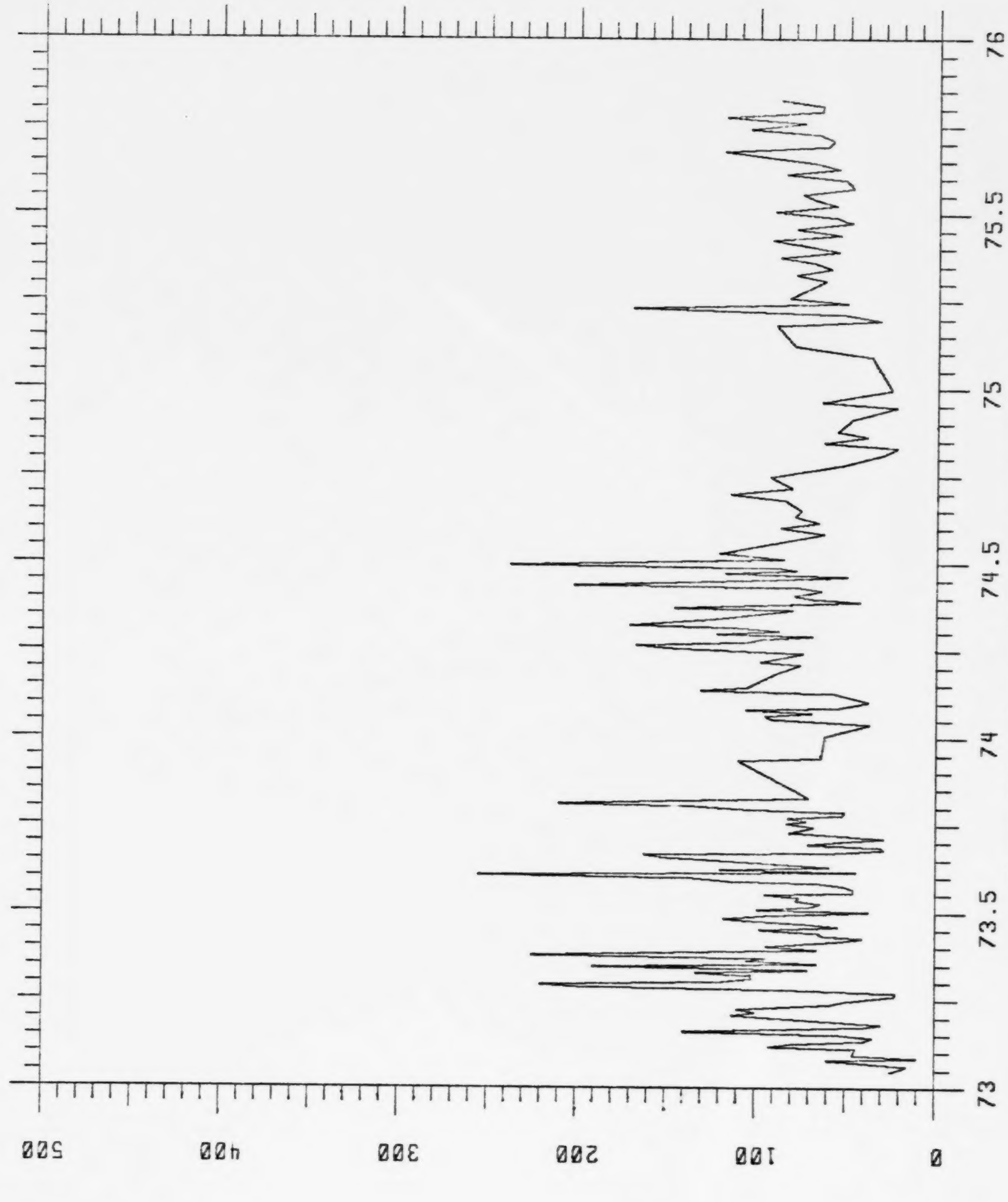


Figure D12 - Time series plot for Bethany, OK

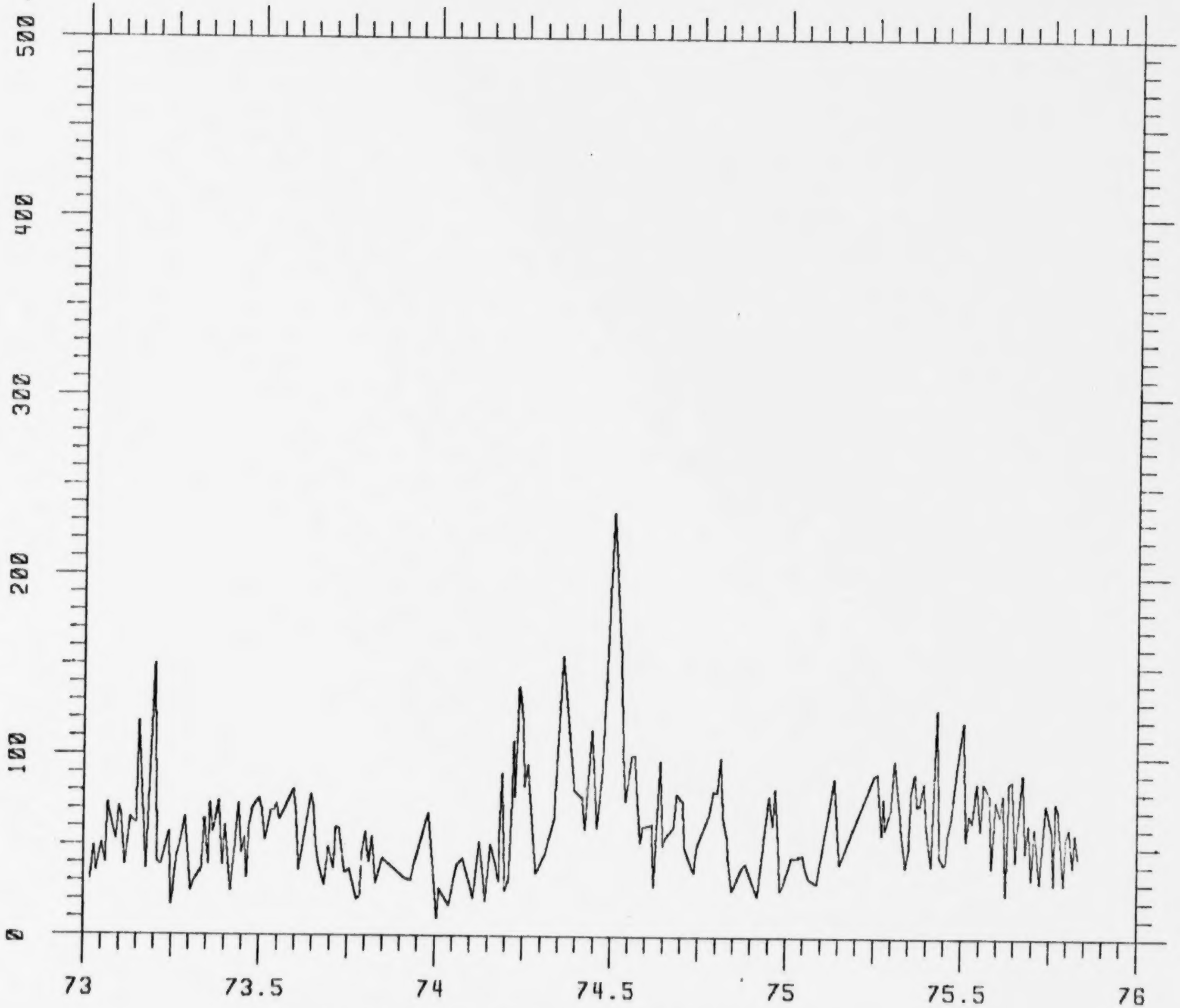


Figure D13 - Time series plot for Muskogee, OK

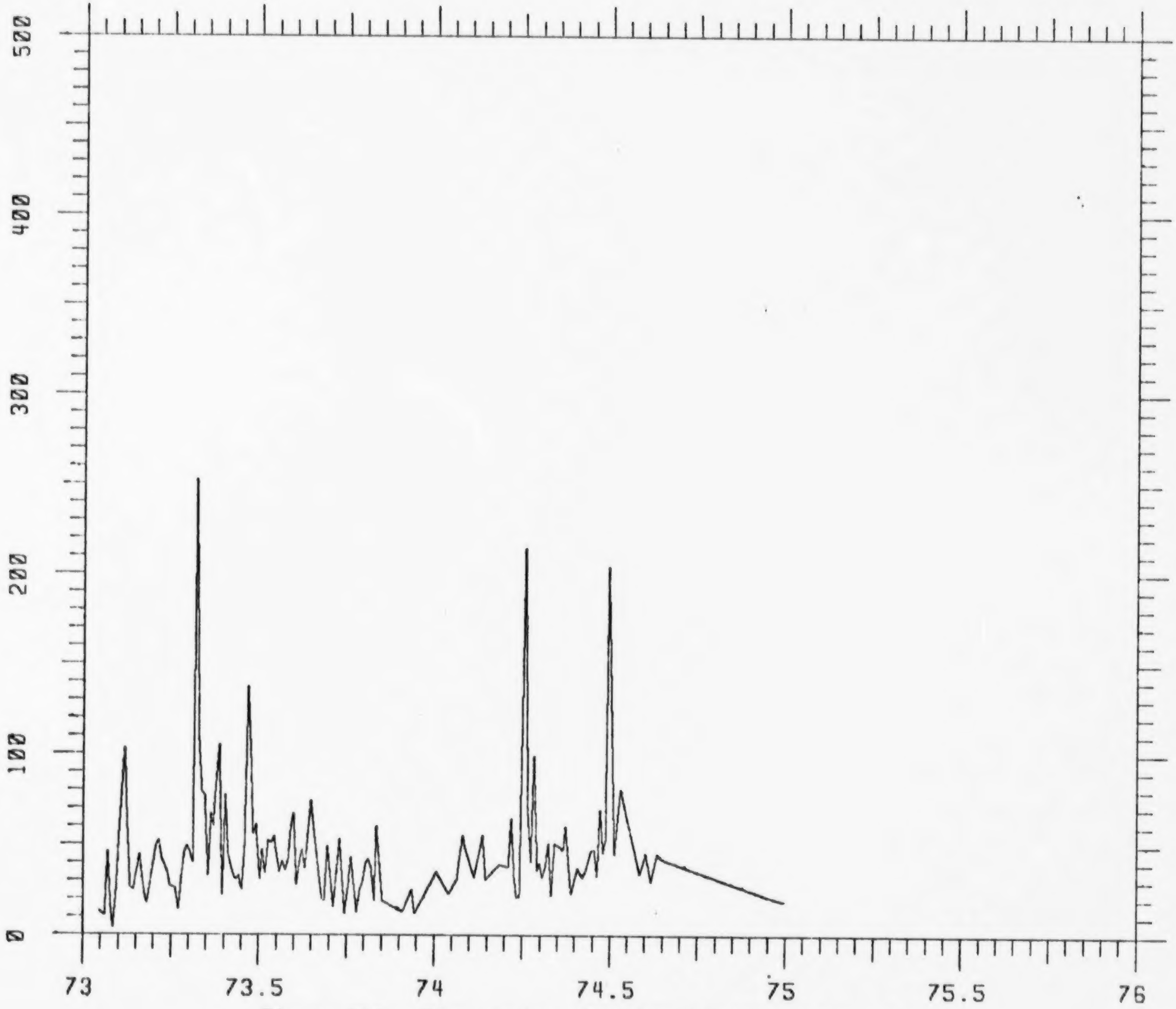


Figure D14 - Time series plot for Oklahoma City, OK

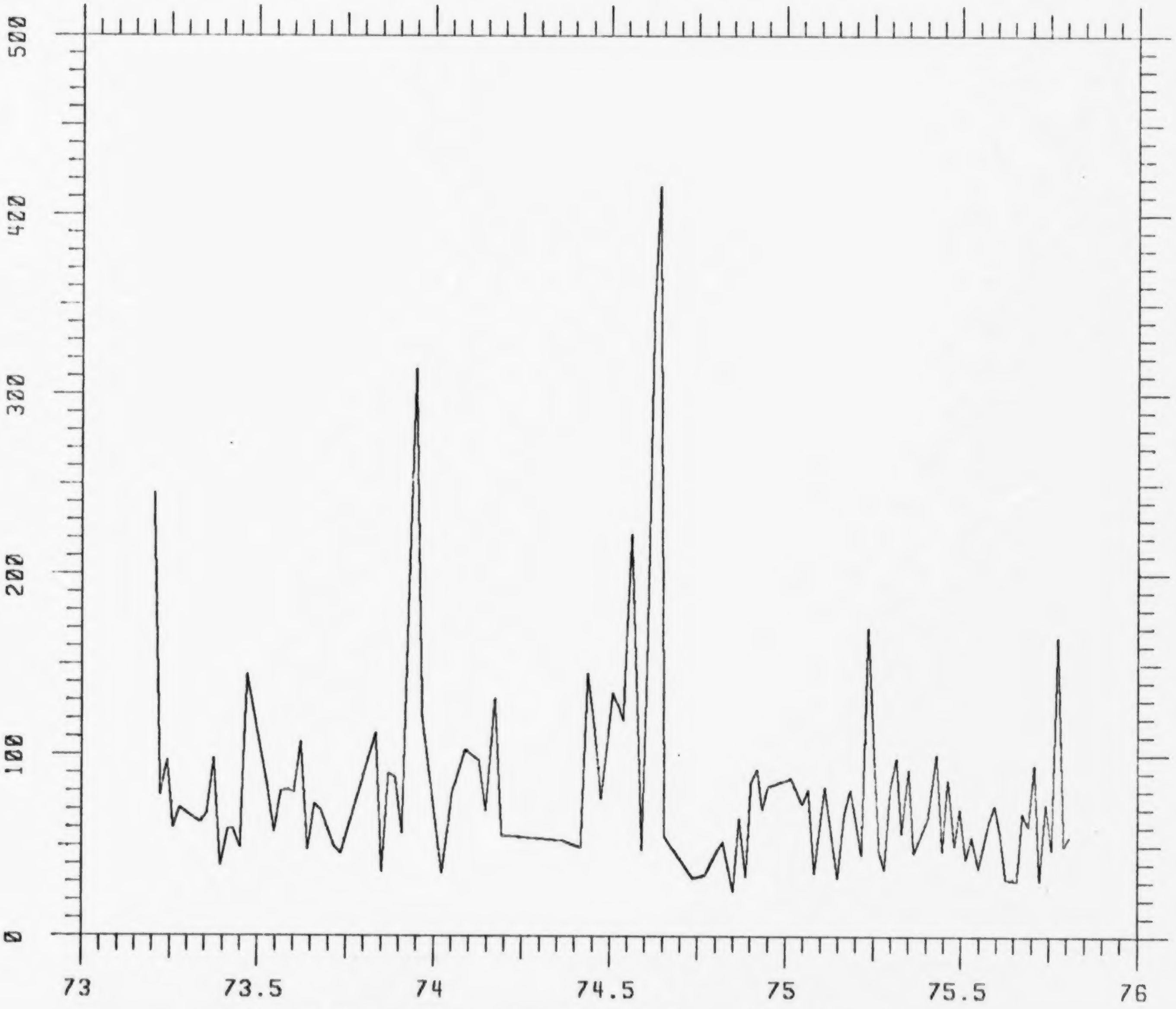


Figure D15 - Time series plot for Roger Mills Co., OK

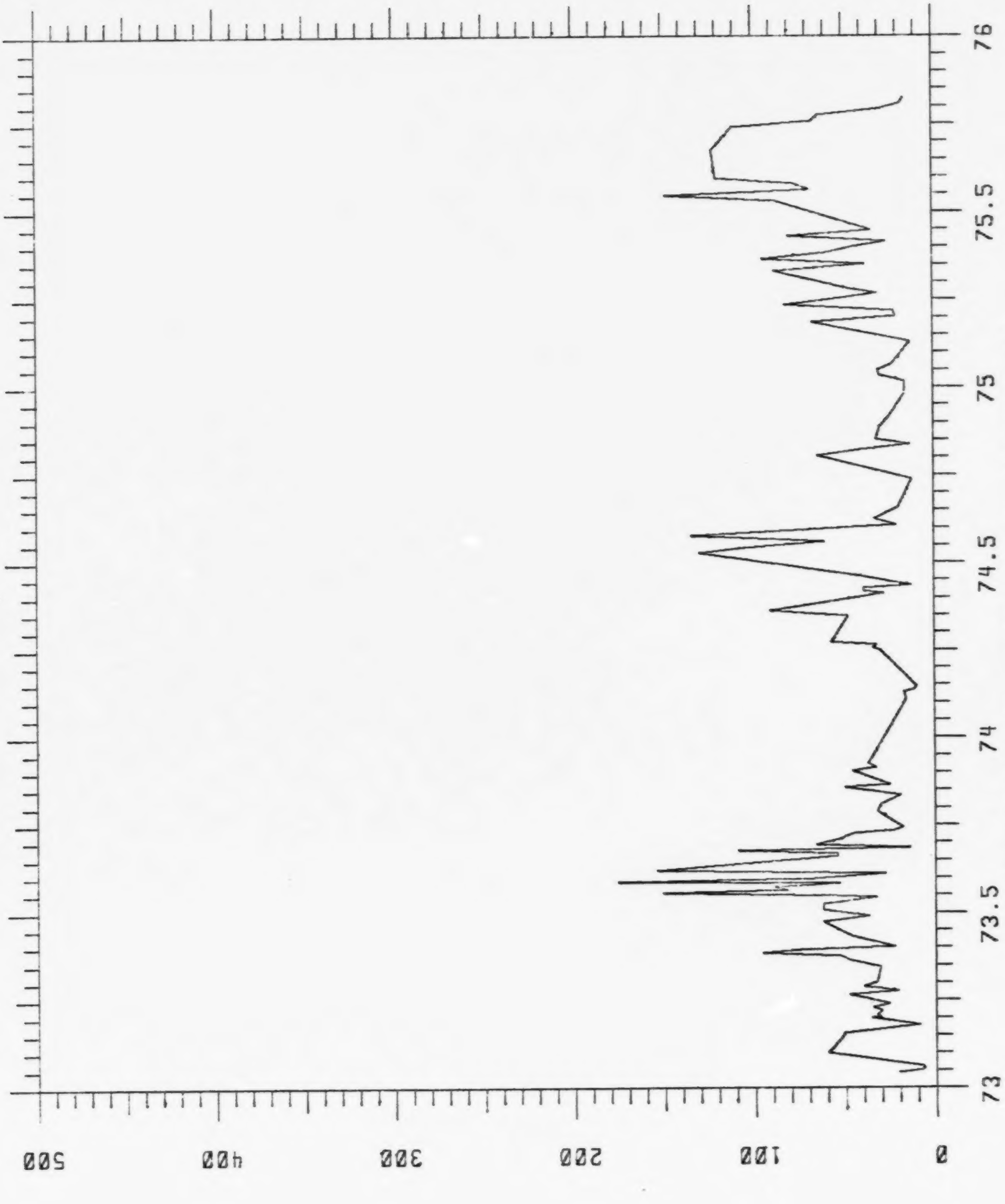


Figure D16 - Time series plot for Sequoyah, OK

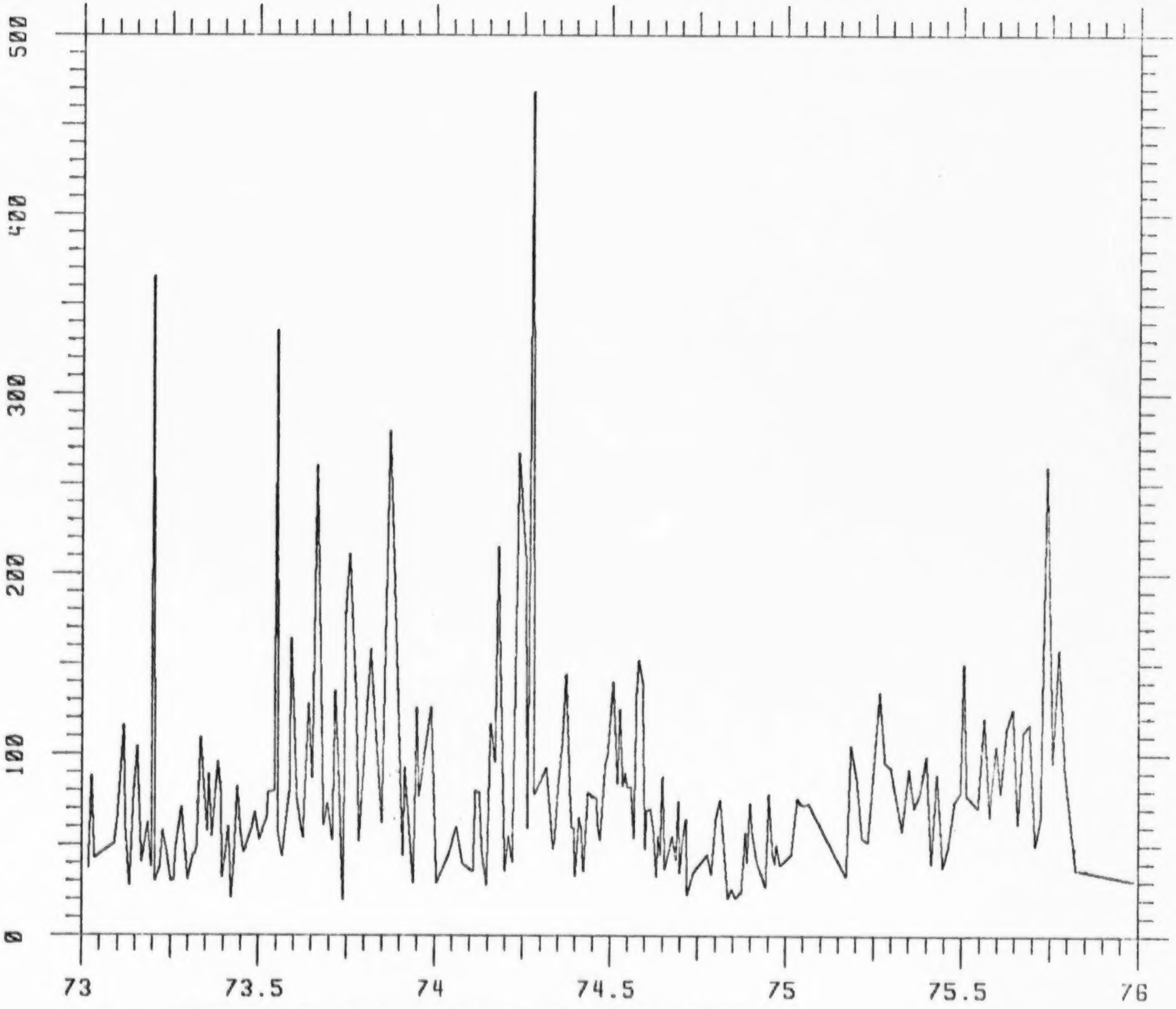


Figure D17 - Time series plot for Tulsa (#1), OK

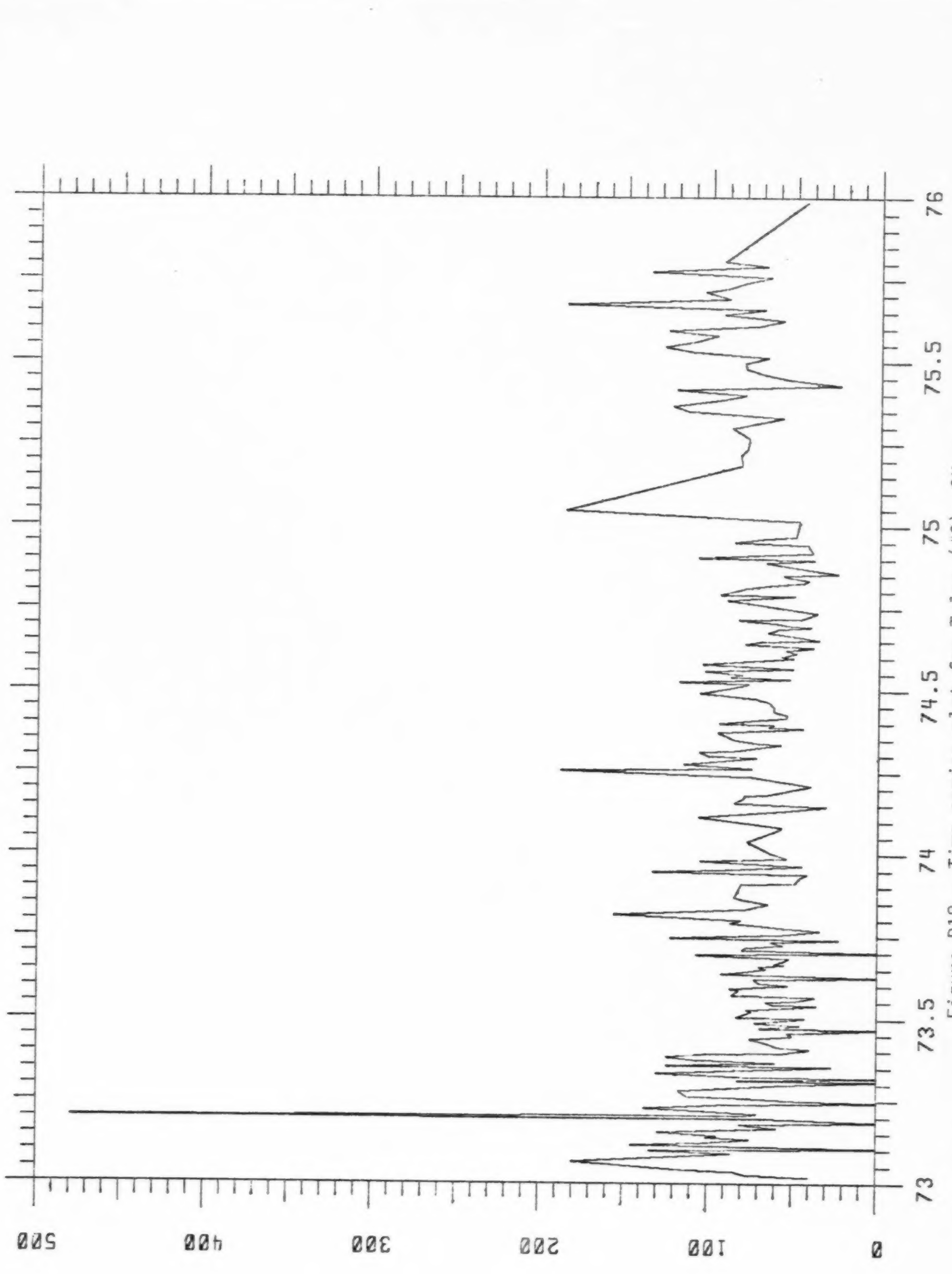


Figure D18 - Time series plot for Tulsa (#2), OK

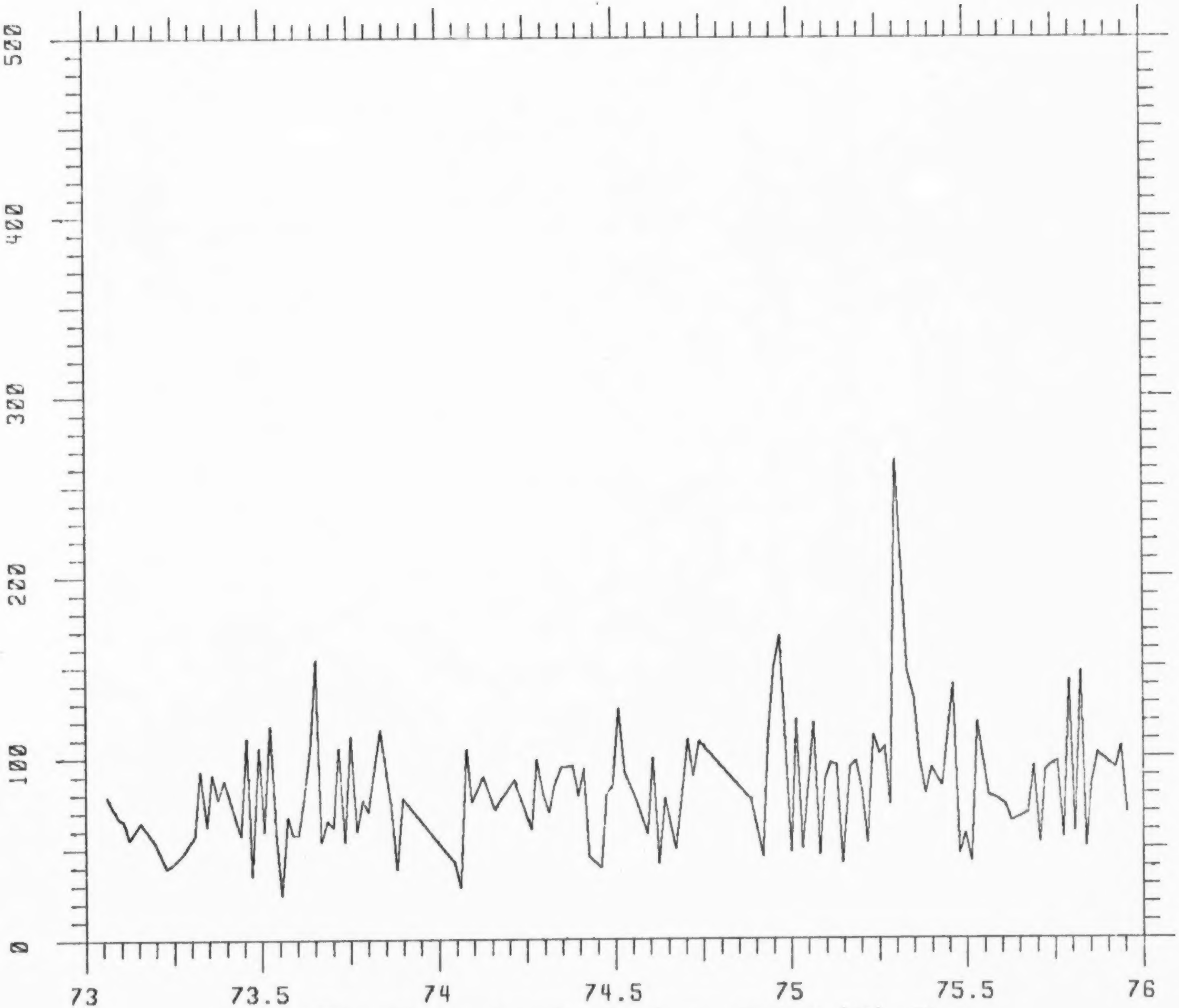


Figure D19 - Time series plot for Corpus Christi (#1), TX

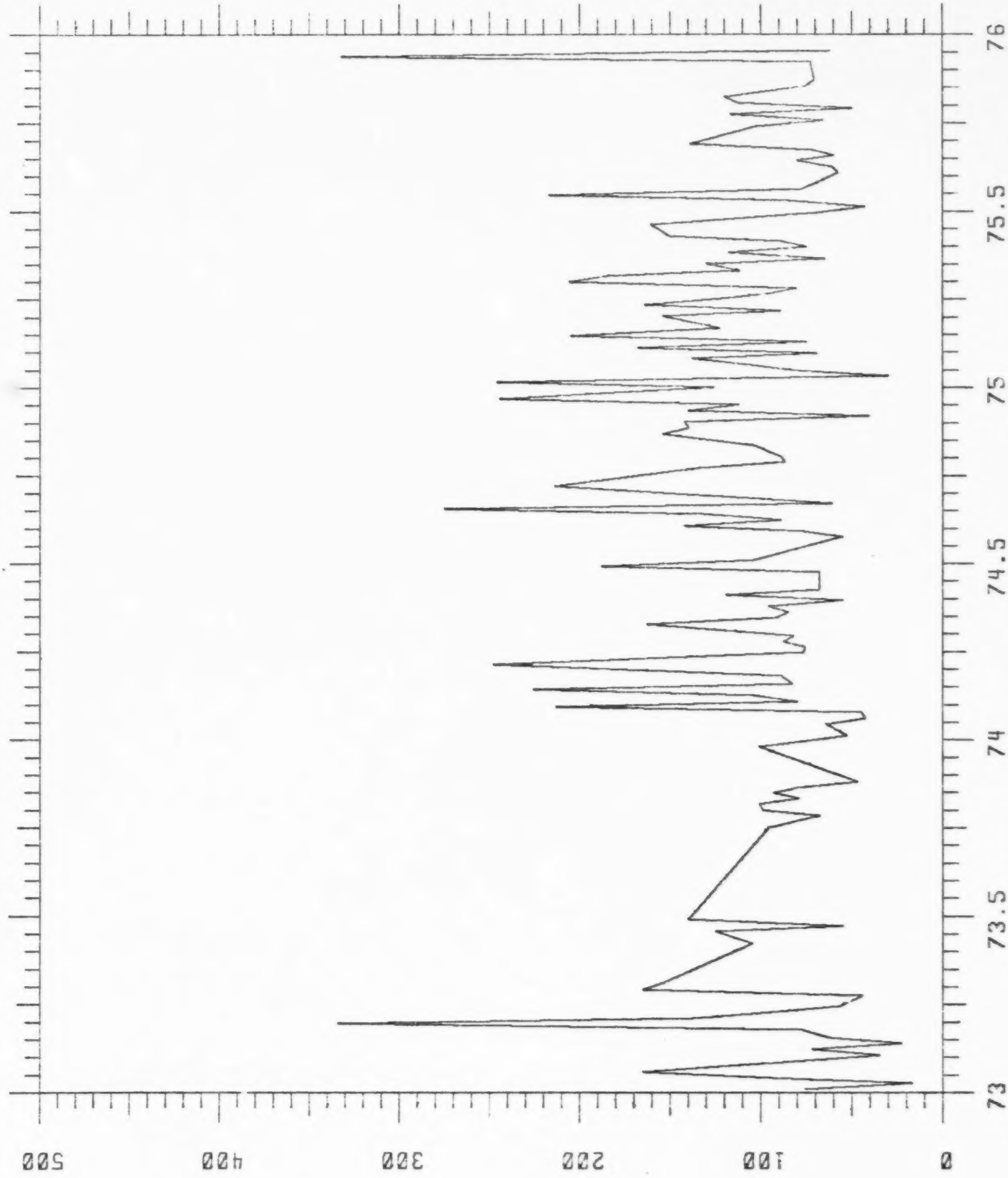


Figure D20 - Time series plot for Corpus Christi (#2), TX

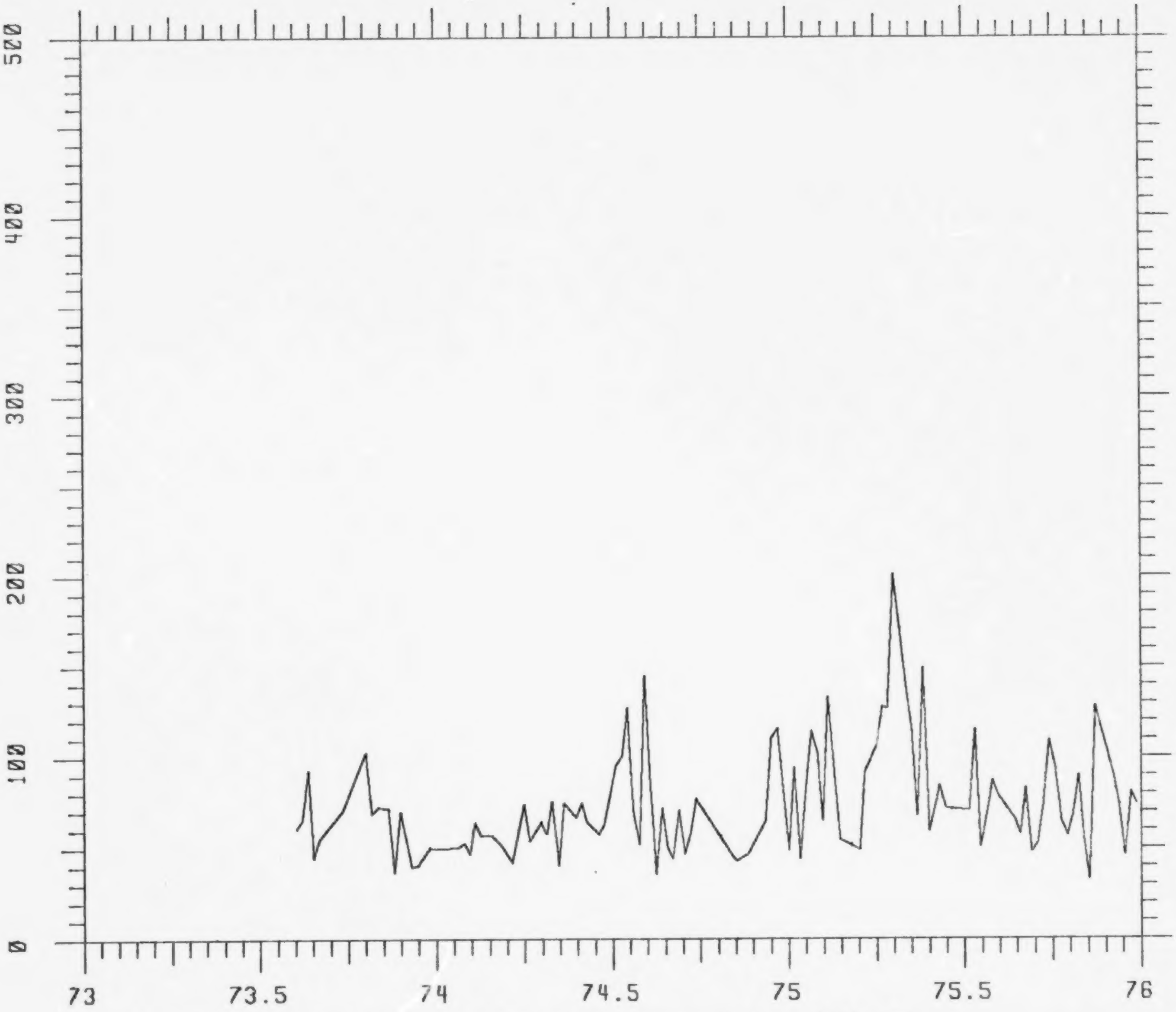


Figure D21 - Time series plot for Corpus Christi (#3), TX

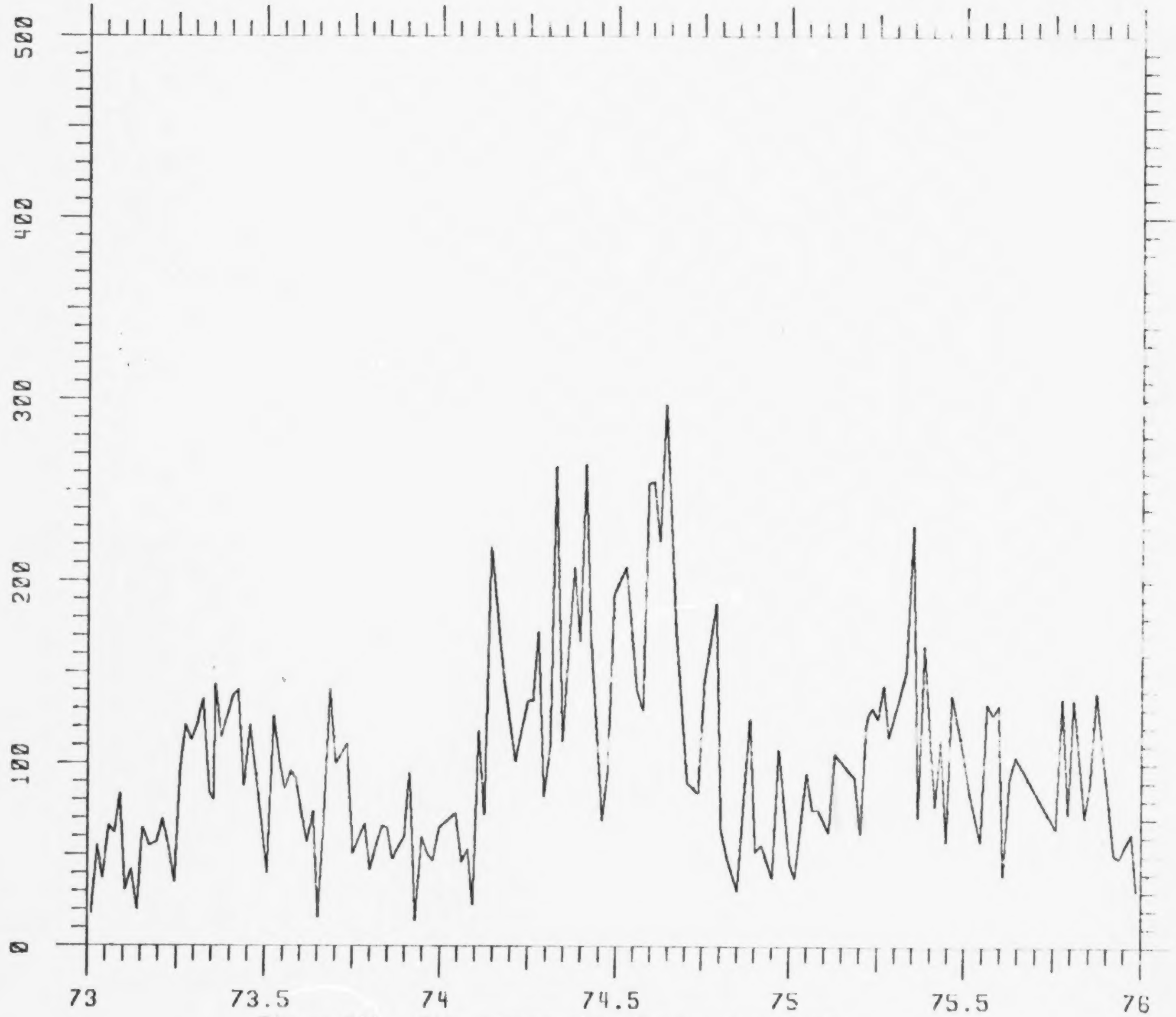


Figure D22 - Time series plot for Harlingen, TX

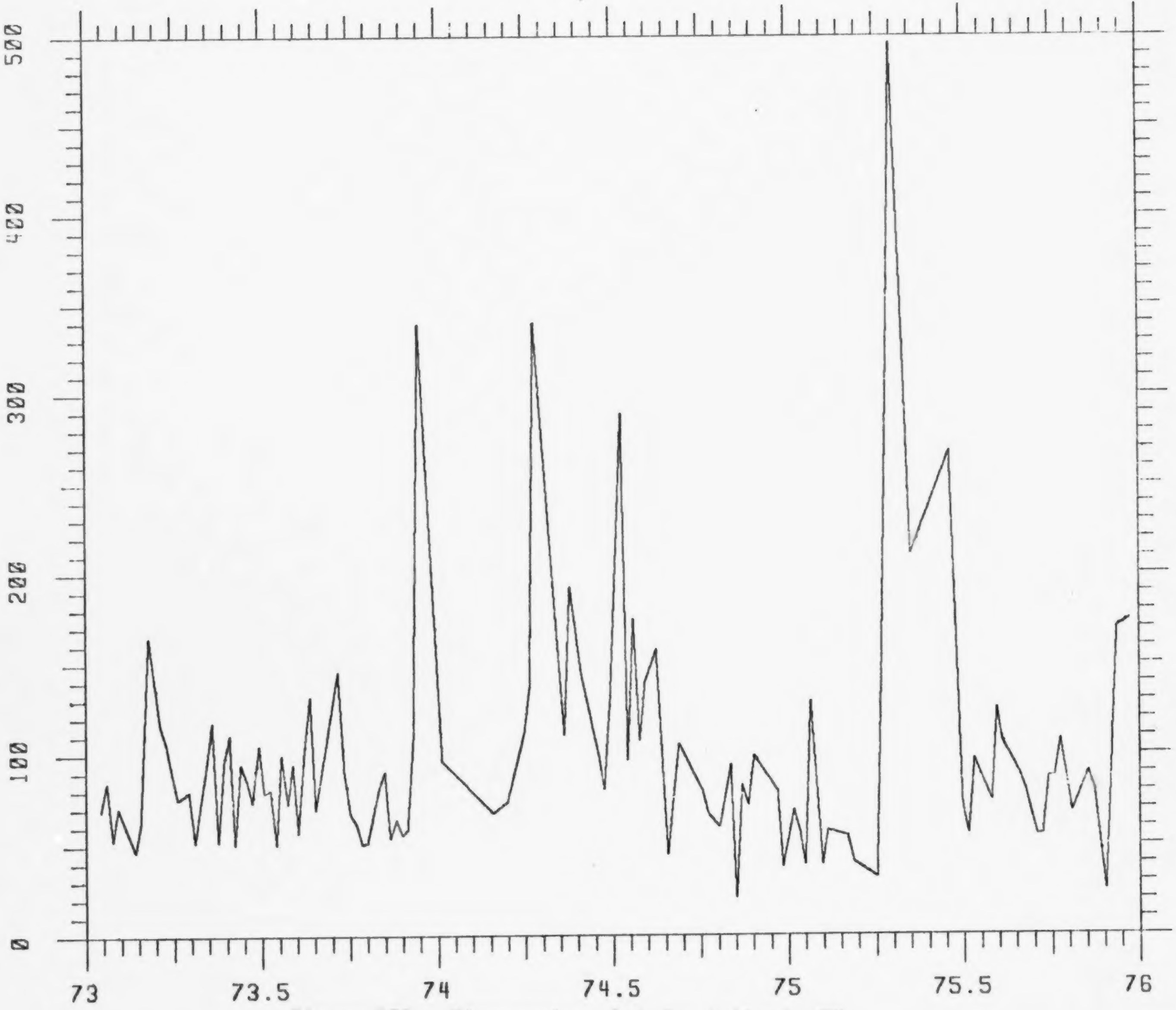


Figure D23 - Time series plot for Lubbock, TX

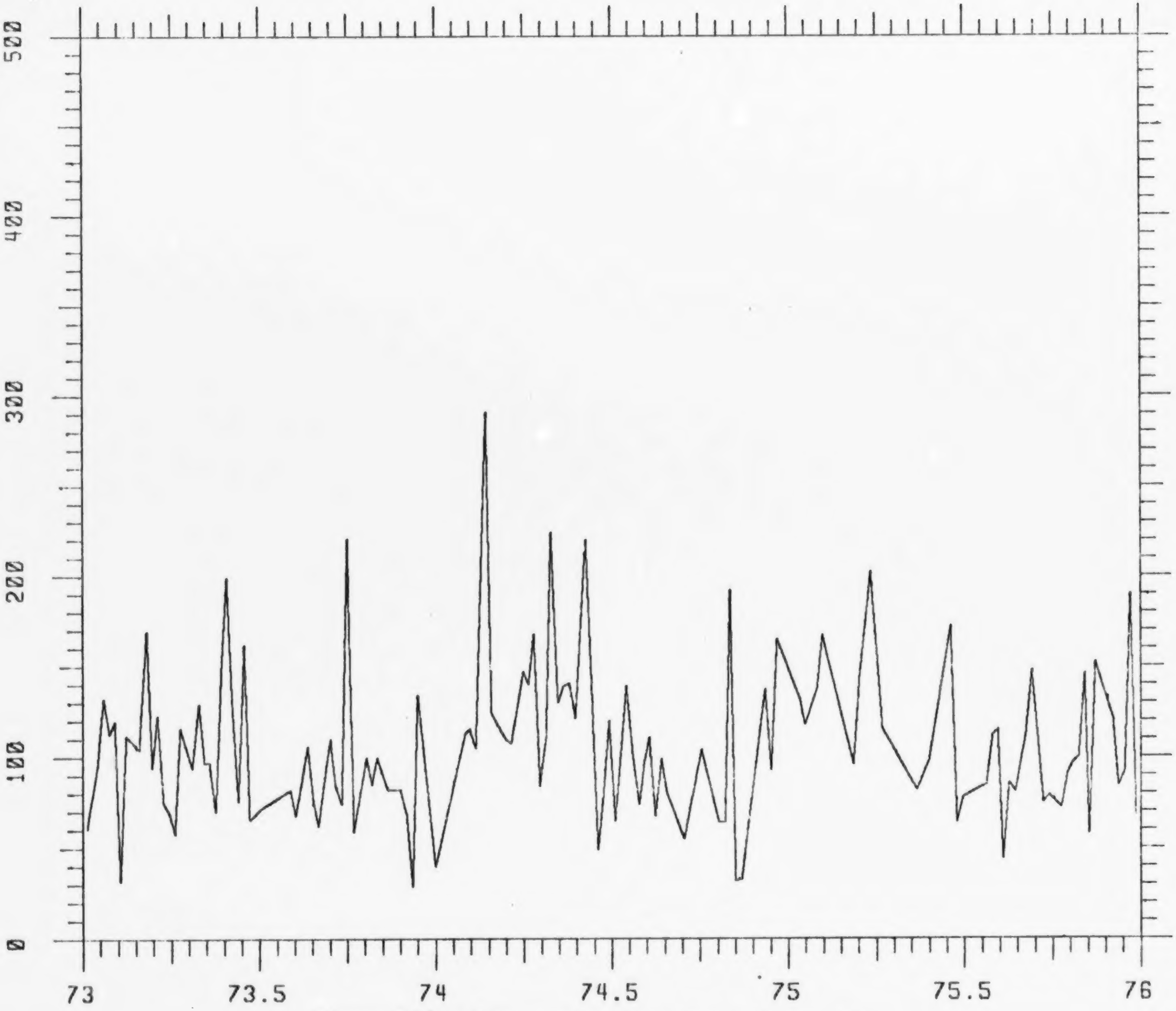


Figure D24 - Time series plot for McAllen, TX

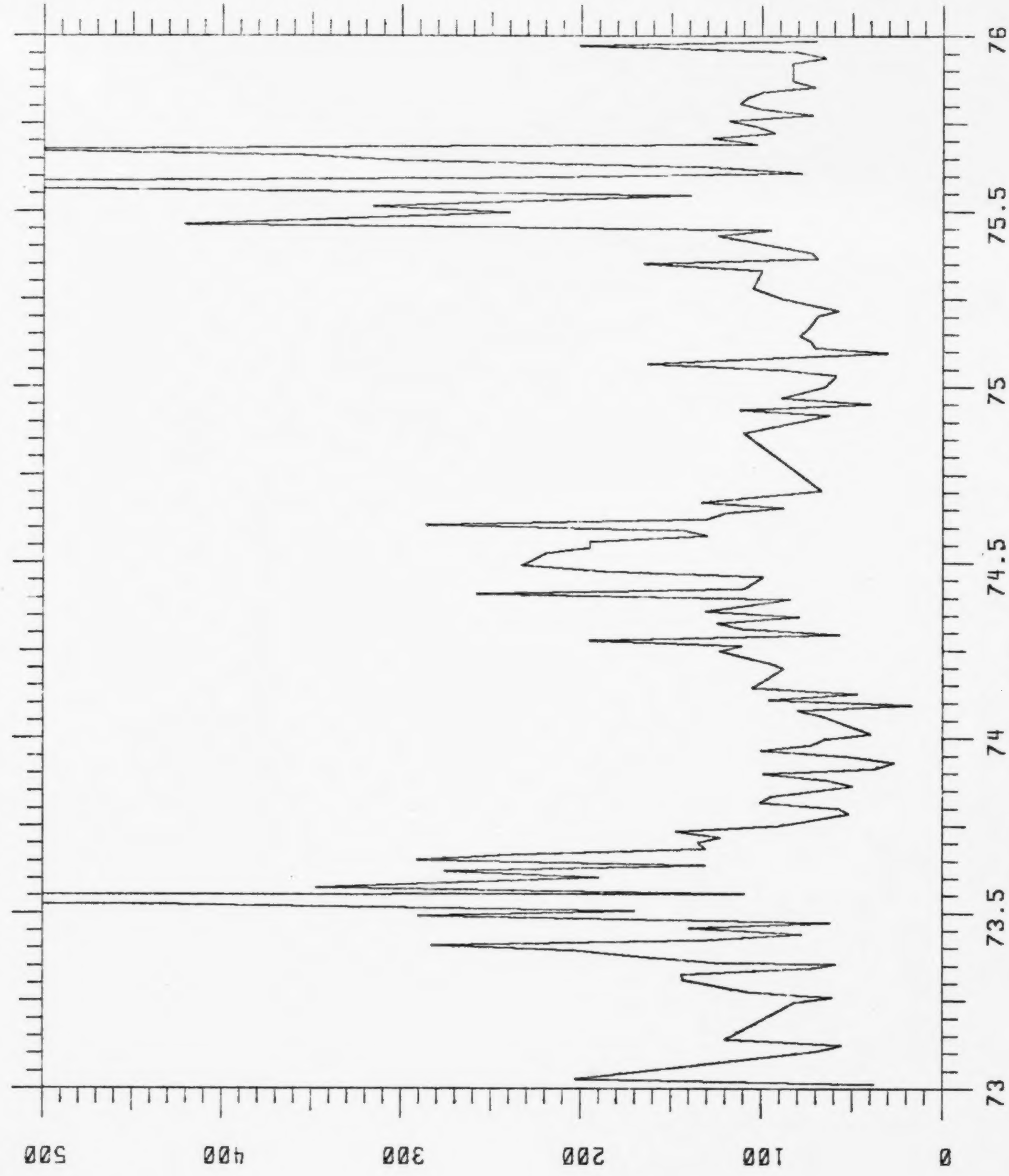


Figure D25 - Time series plot for San Benito, TX